

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 11:52:56 ; Search time 219.652 Seconds
(without alignments)
646.813 Million cell updates/sec

Title: US-10-041-030-8
Perfect score: 24
Sequence: 1 ccagtagtttagcttctgtgctt 24

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	24	6	ABQ78314
2	24	100.0	259	8	Abz18999 Group III
3	24	100.0	554	6	Abt10524 Human bre
4	24	100.0	864	4	Aah34231 Human col
5	24	100.0	3614	4	Aal04355 Human rep
6	24	100.0	3615	4	Aal04354 Human rep
7	24	100.0	3752	4	Aal04353 Human rep
8	24	100.0	5579	8	Acc42349 Human MAP
9	19.2	80.0	2202	9	Adb70333 Thyroid t
10	19.2	80.0	2202	12	Adj37112 Human mal
11	19.2	80.0	3293	4	Aat05810 Human thy
12	19.2	80.0	3293	2	Aas11829 Human TTF
13	18.4	76.7	473	13	Adq56478 Novel can
14	18.2	75.8	621	12	Adk16180 Nanoarcha
15	18.2	75.8	110000	12	Adk16049 Nanoarcha
16	18.2	75.8	143239	12	Adq17729 Human sof
17	17.8	74.2	452	3	Aaa43756 Human sec
18	17.6	73.3	474	6	Abt03344 Ovary cel
19	17.6	73.3	549	6	Abk95242 Human pro
20	17.6	73.3	262090	12	Adq59207 MSI-H car

C	21	17.4	72.5	504	9	ACH38322	Ach38322 Human end
C	22	17.4	72.5	2366	8	ABX10385	Abx10385 DNA encod
C	23	17.4	72.5	2366	13	ADR25645	Adr25645 Breast ca
	24	17.2	71.7	1512	6	ABL42100	Abl42100 Nucleotid
	25	17.2	71.7	15275	4	AAS35975	Aas35975 Human car
	26	17.2	71.7	15275	10	ADE46669	Ade46669 Human car
	27	17.2	71.7	15275	13	ADJ08087	Adj08087 Human car
	28	17.2	71.7	106344	10	ADJ79961	Adj79961 Full leng
	29	17	70.8	101786	3	Aaf22293	Aaf22293 BAC conta
	30	16.8	70.0	514	10	ADC22293	Ade22293 Human lip
C	31	16.8	70.0	564	6	ABQ55470	Abq55470 Human ova
C	32	16.8	70.0	790	2	AAV20469	Aav20469 Human c-r
	33	16.8	70.0	807	6	ABZ42929	Abz42929 Human GPC
	34	16.8	70.0	927	4	AAH31701	Aah31701 Human olf
	35	16.8	70.0	927	12	ADG76823	Adg76823 Human olf
	36	16.8	70.0	958	4	AAH31707	Aah31707 Human olf
	37	16.8	70.0	966	6	ABT04196	Abt04196 Human G-p
	38	16.8	70.0	966	12	ADH30931	Adh30931 Human G-p
	39	16.8	70.0	978	5	AAS42325	Aas42325 Human GPC
	40	16.8	70.0	978	6	ABZ43023	Abz43023 Human GPC
	41	16.8	70.0	978	6	ABK68498	Abk68498 Human DNA
	42	16.8	70.0	978	6	ABK37611	Abk37611 DNA encod
	43	16.8	70.0	978	12	ADG83373	Adg83373 Human Olf
	44	16.8	70.0	1014	8	ABZ58237	Abz58237 Human G-p
	45	16.8	70.0	1330	10	ADC85636	Adc85636 Human GPC

ALIGNMENTS

RESULT 1

ABQ78314
ID ABQ78314 standard; DNA; 24 BP.

AC ABQ78314;

XX 05-NOV-2002 (first entry)

XX Probe N63226QR to detect Pellino 2 gene in cancer cells.

XX Pellino 2; cancer; Pellino 1; cancer treatment; epithelial cancer;
XX gastrointestinal tract cancer; probe; ss.

XX Homo sapiens.

XX WO200259611-A2.

XX 01-AUG-2002.

XX 28-DEC-2003; 2001WO-US051368.

XX 02-JAN-2001; 2001US-0259502P.

XX (TUL) TULARIK INC.

XX Powers S, Mu D, Xiang P, Peng Y;

XX WPI; 2002-619185/66.

XX Detecting cancer cells in mammalian sample, useful for identifying
XX inhibitors for treating cancer e.g. epithelial cancer, comprises
XX detecting an overexpression of, or increase in copy number of genes
XX encoding, Pellino 1 and Pellino 2.

XX Example 2; Page 55; 69pp; English.

XX Probes ABQ78313-15 were used to detect human Pellino 2 gene in cancer
XX cells. The specification describes a method for detecting cancer cells in
XX biological sample from a mammal. The method comprises detecting an
XX overexpression of, or increase in copy number of genes encoding,
XX polypeptides Pellino 1 or Pellino 2. The method is useful in detecting
XX cancer or propensity to develop cancer, monitoring the efficacy of cancer
XX treatment, identifying inhibitors of Pellino 1 and 2, inhibiting the

CC expression and/or activity of Pellino 1 and 2 in cancer cells, and
 CC treating cancer or inhibiting proliferation of cancer. The cancer can be
 CC epithelial cancer, such as lung, colon, ovarian, breast, prostate,
 CC kidney, stomach, bladder, or any cancer of the gastrointestinal tract
 XX
 SQ Sequence 24 BP; 3 A; 5 C; 6 G; 10 T; 0 U; 0 Other;
 Query Match 100.0%; Score 24; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.18; Mismatches 0; Indels 0; Gaps 0;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAGTAGTTAGCCTTTGGGCTT 24
 DB 1 CCAGTAGTTAGCCTTTGGGCTT 24
 RESULT 2
 ABZ18999/c
 ID ABZ18999 standard; cDNA; 259 BP.
 XX
 AC ABZ18999;
 XX
 XX
 DT 23-JAN-2003 (first entry)
 XX
 DE Group III cDNA cancer related clone SEQ ID NO:1425.
 XX
 DE Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200278516-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US010421.
 XX
 XX 30-MAR-2001; 2001US-0280255P.
 PR 28-AUG-2001; 2001US-0315563P.
 PR 09-JAN-2002; 2002US-0347313P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Wang T, Wang S, Bangur CS, Gaiger A;
 PI
 XX WPI; 2003-058387/05.
 DR
 XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
 PT preventing and treating cancer, expressing CT or CP mRNA antigens, and in
 PT virology, immunology, microbiology, molecular biology and recombinant DNA
 PT techniques.
 XX
 PS Claim 1; SEQ ID NO 1425; 207pp; English.
 XX
 CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
 CC invention. (I) and (II) have cytostatic activity and can be used in gene
 CC therapy and vaccines. (I), (II), antibodies and compositions from the
 CC present invention are useful for diagnosing, preventing and treating
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for
 CC stimulating immune response. They can also be useful in virology,
 CC immunology, microbiology, molecular biology and recombinant DNA
 CC techniques. N.B. The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 259 BP; 89 A; 50 C; 59 G; 61 T; 0 U; 0 Other;
 Query Match 100.0%; Score 24; DB 8; Length 259;
 Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGGCTT 24
 DB 79 CCAGTAGTTAGCCTTTGGGCTT 56
 RESULT 3
 ABT10524/c
 ID ABT10524 standard; cDNA; 554 BP.
 XX
 AC ABT10524;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE Human breast cancer associated coding sequence SEQ ID NO: 658.
 XX
 DE Human; breast specific gene; breast cancer; differential expression;
 KW cytostatic; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200259271-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 25-JAN-2002; 2002WO-US002176.
 XX
 XX 25-JAN-2001; 2001US-0263757P.
 PR 25-APR-2001; 2001US-0286090P.
 PR 23-MAY-2001; 2001US-0292517P.
 XX
 XX (GENE-) GENE LOGIC INC.
 PA
 XX Orr MS, Nation M, Diggans JC, Zeng W;
 PI
 XX WPI; 2002-674803/72.
 DR
 XX Diagnosing breast cancer in a patient comprises detecting the level of
 PT gene expression in cell or tissue samples, where a differential gene
 PT expression is indicative of breast cancer.
 XX
 PS Claim 1; SEQ ID NO 658; 260pp + Sequence Listing; English.
 XX
 CC The present invention relates to methods of diagnosing breast cancer in a
 CC patient, which comprise detecting the level of expression in a tissue
 CC sample of two or more genes selected from those shown in ABT09867.
 CC ABT1112, where a differential expression of the genes indicates breast
 CC cancer. The methods are useful in diagnosing, treating, detecting the
 CC progression, and in monitoring treatment of breast cancer in patients.
 CC The methods are also useful as a screening tool for agents that modulate
 CC the onset or progression of breast cancer. The breast cancer genes may be
 CC used as diagnostic markers for the prediction or identification of the
 CC malignant state of breast tissue, for confirming the type and progression
 CC of cancer, and for drug screening and assays. The present sequence is a
 CC coding sequence of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 554 BP; 196 A; 88 C; 92 G; 178 T; 0 U; 0 Other;
 Query Match 100.0%; Score 24; DB 6; Length 554;
 Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAGTAGTTAGCCTTTGGGCTT 24
 DB 519 CCAGTAGTTAGCCTTTGGGCTT 496
 RESULT 4
 AAH34231
 ID AAH34231 standard; cDNA; 864 BP.
 XX

AC AAH34231;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1313.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN W0200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Baraah SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR P-PSDB; AAG74826.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 1; Page 3049; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37195 to AAH37204 and AAG77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 864 BP; 247 A; 143 C; 157 G; 314 T; 0 U; 3 Other;

Query Match 100.0%; Score 24; DB 4; Length 864;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCAGTAGTTAGCCTTTGGCTT 24
Db 260 CCAGTAGTTAGCCTTTGGCTT 283

RESULT 5
AAL04355
ID AAL04355 standard; DNA; 3614 BP.
XX
AC AAL04355;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7043.
XX
KW Human; reproductive system related antigen; reproductive system disorder;

KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN W0200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.

```
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246471P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250331P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251088P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
PT used in preventing, treating or ameliorating a medical condition.
XX PS Disclosure; SEQ ID NO 7043; 1297pp + Sequence Listing; English.
XX CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX SQ Sequence 3614 BP; 1063 A; 621 C; 657 G; 1273 T; 0 U; 0 Other;
Query Match 100.0%; Score 24; DB 4; Length 3614;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCAGTAGTTTACCTTTGTGGCTT 24
Db 3012 CCAGTAGTTTACCTTTGTGGCTT 3035
RESULT 6
AAL04354
ID AAL04354 standard; DNA; 3615 BP.
XX AC AAL04354;
XX DT 21-NOV-2001 (first entry)
XX DE Human reproductive system related antigen DNA SEQ ID NO: 7042.
XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KW cancer; gene therapy; ds.
XX OS Homo sapiens.
XX PN WO200155320-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001339.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214864P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
```


Thu Jul 21 09:13:22 2005

us-10-041-030-8.rng

AC	AA04353;	PR	14-SEP-2000;	2000US-0232401P.
XX		PR	14-SEP-2000;	2000US-0233063P.
DT	21-NOV-2001 (first entry)	PR	14-SEP-2000;	2000US-0233064P.
XX		PR	14-SEP-2000;	2000US-0233065P.
DE	Human reproductive system related antigen DNA SEQ ID NO: 7041.	PR	21-SEP-2000;	2000US-0234223P.
XX		PR	21-SEP-2000;	2000US-0234274P.
KW	Human; reproductive system related antigen; reproductive system disorder;	PR	25-SEP-2000;	2000US-0234997P.
KW	cancer; gene therapy; ds.	PR	25-SEP-2000;	2000US-0234998P.
XX		PR	26-SEP-2000;	2000US-0235484P.
XX		PR	27-SEP-2000;	2000US-0235834P.
OS	Homo sapiens.	PR	27-SEP-2000;	2000US-0235836P.
XX	WO200155320-A2.	PR	29-SEP-2000;	2000US-0236327P.
PN		PR	29-SEP-2000;	2000US-0236367P.
XX		PR	29-SEP-2000;	2000US-0236368P.
PD		PR	29-SEP-2000;	2000US-0236369P.
XX		PR	29-SEP-2000;	2000US-0236370P.
XX		PR	02-OCT-2000;	2000US-0236802P.
PF	17-JAN-2001; 2001WO-US001339.	PR	02-OCT-2000;	2000US-0237037P.
XX		PR	02-OCT-2000;	2000US-0237038P.
XX	31-JAN-2000; 2000US-0179065P.	PR	02-OCT-2000;	2000US-0237039P.
PR	04-FEB-2000; 2000US-0180628P.	PR	02-OCT-2000;	2000US-0237040P.
PR	24-FEB-2000; 2000US-0184664P.	PR	02-OCT-2000;	2000US-0239935P.
PR	02-MAR-2000; 2000US-0186350P.	PR	13-OCT-2000;	2000US-0239937P.
PR	16-MAR-2000; 2000US-0189874P.	PR	20-OCT-2000;	2000US-0240960P.
PR	17-MAR-2000; 2000US-0190076P.	PR	20-OCT-2000;	2000US-0241221P.
PR	18-APR-2000; 2000US-0198123P.	PR	20-OCT-2000;	2000US-0241785P.
PR	19-MAY-2000; 2000US-0205515P.	PR	20-OCT-2000;	2000US-0241786P.
PR	07-JUN-2000; 2000US-0209467P.	PR	20-OCT-2000;	2000US-0241808P.
PR	28-JUN-2000; 2000US-0214886P.	PR	20-OCT-2000;	2000US-0241809P.
PR	30-JUN-2000; 2000US-0215135P.	PR	20-OCT-2000;	2000US-0241826P.
PR	07-JUL-2000; 2000US-0216647P.	PR	20-OCT-2000;	2000US-0244617P.
PR	07-JUL-2000; 2000US-0216880P.	PR	08-NOV-2000;	2000US-0246474P.
PR	11-JUL-2000; 2000US-0217487P.	PR	08-NOV-2000;	2000US-0246475P.
PR	11-JUL-2000; 2000US-0217496P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-JUL-2000; 2000US-0218290P.	PR	08-NOV-2000;	2000US-0246477P.
PR	26-JUL-2000; 2000US-0220963P.	PR	08-NOV-2000;	2000US-0246478P.
PR	26-JUL-2000; 2000US-0220964P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000; 2000US-0224518P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000; 2000US-0224519P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000; 2000US-0225266P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000; 2000US-0225271P.	PR	08-NOV-2000;	2000US-0246611P.
PR	14-AUG-2000; 2000US-0225275P.	PR	08-NOV-2000;	2000US-0246613P.
PR	14-AUG-2000; 2000US-0225275P.	PR	08-NOV-2000;	2000US-0249207P.
PR	14-AUG-2000; 2000US-0225275P.	PR	17-NOV-2000;	2000US-0249208P.
PR	14-AUG-2000; 2000US-0226681P.	PR	17-NOV-2000;	2000US-0249209P.
PR	22-AUG-2000; 2000US-0226686P.	PR	17-NOV-2000;	2000US-0249210P.
PR	22-AUG-2000; 2000US-0227182P.	PR	17-NOV-2000;	2000US-0249211P.
PR	22-AUG-2000; 2000US-0227009P.	PR	17-NOV-2000;	2000US-0249212P.
PR	30-AUG-2000; 2000US-0228924P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000; 2000US-0229287P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000; 2000US-0229343P.	PR	17-NOV-2000;	2000US-0249215P.
PR	01-SEP-2000; 2000US-0229344P.	PR	17-NOV-2000;	2000US-0249216P.
PR	01-SEP-2000; 2000US-0229345P.	PR	17-NOV-2000;	2000US-0249217P.
PR	05-SEP-2000; 2000US-0229509P.	PR	17-NOV-2000;	2000US-0249218P.
PR	05-SEP-2000; 2000US-0229513P.	PR	17-NOV-2000;	2000US-0249219P.
PR	06-SEP-2000; 2000US-0230437P.	PR	17-NOV-2000;	2000US-0249224P.
PR	06-SEP-2000; 2000US-0230438P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000; 2000US-0231242P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000; 2000US-0231243P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000; 2000US-0231244P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000; 2000US-0231413P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000; 2000US-0232080P.	PR	17-NOV-2000;	2000US-0249224P.
PR	08-SEP-2000; 2000US-0232081P.	PR	17-NOV-2000;	2000US-0249224P.
PR	12-SEP-2000; 2000US-0232081P.	PR	17-NOV-2000;	2000US-0249224P.
PR	14-SEP-2000; 2000US-0232337P.	PR	17-NOV-2000;	2000US-0249224P.
PR	14-SEP-2000; 2000US-0232338P.	PR	17-NOV-2000;	2000US-0249224P.
PR	14-SEP-2000; 2000US-0232339P.	PR	17-NOV-2000;	2000US-0249224P.
PR	14-SEP-2000; 2000US-0232339P.	PR	17-NOV-2000;	2000US-0249224P.
PR	14-SEP-2000; 2000US-0232400P.	PR	17-NOV-2000;	2000US-0249224P.

```
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Baraah SC, Ruben SM;
XX
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 7041; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX
XX Sequence 3752 BP; 1101 A; 648 C; 671 G; 1332 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 24; DB 4; Length 3752;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCAGTAGTTAGCCTTTGGCTT 24
DB 3013 CCAGTAGTTAGCCTTTGGCTT 3036
RESULT 8
ACC42349
ID ACC42349 standard; cDNA; 5579 BP.
XX
XX ACC42349;
XX
XX 22-MAY-2003 (first entry)
XX
XX Human MAP kinase cascade activator #59 cDNA.
XX
XX Human; Elkl phosphorylation; Elkl phosphorylation kinase; virucide;
XX antiinflammatory; immunomodulator; cytostatic; antiallergic; anti-HIV;
XX antirheumatic; antiarthritic; antidiabetic; antiasthmatic; gene therapy;
XX inflammation; autoimmune disease; viral disease; cancer; diabetes;
XX rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
XX Iga nephritis; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003008589-A1.
XX
XX 30-JAN-2003.
XX
XX 15-JUL-2002; 2002WO-JP007174.
XX
XX 18-JUL-2001; 2001JP-00218204.
XX 31-AUG-2001; 2001JP-00263450.
XX 21-JAN-2002; 2002JP-00012176.
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
XX
XX WPI; 2003-229582/22.
XX P-PSDB; ABR41083.
XX
XX Elkl phosphorylation-associated gene and its encoded protein with MAP
PT
kinase cascade effect, applicable in diagnosis of and developing drugs
for e.g. inflammations, autoimmune diseases, viral diseases and cancer.
XX
XX Claim 4; Page 631-640; 762pp; Japanese.
XX
XX The invention relates to a novel purified protein having Elkl
phosphorylation activity and/or an activity of activating Elkl
phosphorylation kinase. A protein of the invention has antinflammatory,
immunomodulator, virucide, cytostatic, antiallergic, antirheumatic,
antiarthritic, antidiabetic, antiasthmatic, and anti-HIV activity. The
polynucleotides may have a use in gene therapy. The gene and its encoded
protein are applicable in diagnosis of and developing drugs for e.g.
inflammations, autoimmune diseases, viral diseases and cancer such as
rheumatoid arthritis, diabetes; asthma, allergic rhinitis, AIDS, viral
hepatitis and Iga nephritis. The present sequence is used in the
exemplification of the invention
XX
XX Sequence 5579 BP; 1535 A; 1111 C; 1214 G; 1719 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 24; DB 8; Length 5579;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCAGTAGTTAGCCTTTGGCTT 24
DB 5062 CCAGTAGTTAGCCTTTGGCTT 5085
RESULT 9
ADB70333/c
ID ADB70333 standard; cDNA; 2202 BP.
XX
XX ADB70333;
XX
XX 04-DEC-2003 (first entry)
XX
XX Thyroid transcription factor 1 cDNA SEQ ID NO:25.
XX
XX cancer; malignant pleural mesothelioma; MPN; lung adenocarcinoma;
XX squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
XX diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
XX human; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003021229-A2.
XX
XX 13-MAR-2003.
XX
XX 05-SEP-2002; 2002WO-US028203.
XX
XX 05-SEP-2001; 2001US-0317389P.
XX 30-AUG-2002; 2002US-00236031.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Gordon GJ, Jensen RV, Gullans SR, Bueno R;
XX
XX WPI; 2003-290233/28.
XX P-PSDB; ADB70370.
XX
XX Diagnosing cancer cells in tissue sample, or determining prognosis or
outcome of cancer patient, by calculating ratio of expression levels of
genes that are differentially expressed in cancer and non cancer tissues.
XX
XX Claim 67; Page 141-142; 396pp; English.
XX
XX The present invention describes a method (M1) for diagnosing the presence
of cancer cells or non-cancer cells in a tissue sample, or determining
the prognosis or outcome of a cancer patient. M1 involves providing a set
of genes that are differentially expressed in cancerous or non-cancerous
conditions, determining the expression levels of the set of genes and
calculating a ratio of the expression levels of the differentially
```

expressed genes. M1 is useful for diagnosing the presence of cancer cells or non-cancer cells in a tissue sample, where the cancer is malignant. CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma, CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell lymphoma, follicular lymphoma and ovarian cancer, and for determining CC prognosis or outcome of a cancer patient. The ratio of expression levels of differentially expressed genes is used as an indicator of cancer type, CC cancer class, and/or cancer prognosis, all of which are useful for CC determining a course of treatment of a patient. The present sequence CC encodes a human protein which is used in an example from the present CC invention.

Query Match 80.0%; Score 19.2; DB 9; Length 2202;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCAGTAGTTAGCCTTTGGCTT 24
Db 1451 CCAGCAGTTGGCCTTTGGCTT 1428

RESULT 10
ADJ37112/c
ID ADJ37112 standard; cDNA; 2202 BP.

AC ADJ37112;
XX
XX
DT 22-APR-2004 (first entry)
XX
XX Human malignant pleural mesothelioma (MPM) cDNA #9.
XX
XX Human; malignant pleural mesothelioma; MPM; gene; ss; tumour;
KW lung adenocarcinoma; squamous carcinoma; medulloblastoma;
KW prostate cancer; breast cancer; diffuse large B-cell lymphoma;
KW follicular lymphoma; ovarian cancer; cytostatic.

XX Homo sapiens.
XX
XX US2003219760-A1.
XX
XX 27-NOV-2003.
XX
XX 05-SEP-2002; 2002US-00236031.
XX
XX 05-SEP-2001; 2001US-0317389P.
XX
XX 30-AUG-2002; 2002US-0407431P.
XX
XX (SGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Gordon GJ, Jensen RV, Gullans SR, Bueno R;
XX
XX WPI; 2004-141744/14.
XX
XX P-PSDB; ADJ37113.

Diagnosing the presence of cancer or non-cancer cells in tissue sample,
useful for diagnosing malignant pleural mesothelioma comprises
determining ratio of expression level of a set of genes expressed in
cancer tissues.

Claim 44; SEQ ID NO 25; 53pp; English.
XX
XX The invention relates to a method of diagnosing the presence of cancer
cells or non-cancer cells in a tissue sample, determining prognosis or
outcome of a cancer patient, selecting a course of treatment for a
subject having or suspected of having malignant pleural mesothelioma
(MPM) and evaluating treatment of MPM comprising determining the ratio of
the expression level of a set of genes differentially expressed in a
cancer tissue. The cancer is chosen from MPM, lung adenocarcinoma,
squamous carcinoma, medulloblastoma, prostate cancer, breast cancer,
diffuse large B-cell lymphoma, follicular lymphoma and ovarian cancer.
XX The method is useful for diagnosing MPM in a subject suspected of having

CC MPM which involves obtaining a tissue sample suspected of being cancerous
from a subject and determining the expression of nucleic acid markers or
its expression products in the tissue sample. This sequence represents
CC human MPM cDNA of the invention. Note: The sequence data for this patent
did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

Qy Sequence 2202 BP; 503 A; 662 C; 643 G; 394 T; 0 U; 0 Other;
Query Match 80.0%; Score 19.2; DB 12; Length 2202;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCAGTAGTTAGCCTTTGGCTT 24
Db 1451 CCAGCAGTTGGCCTTTGGCTT 1428

RESULT 11
AAT05810/c
ID AAT05810 standard; DNA; 3293 BP.

AC AAT05810;
XX
XX
DT 27-NOV-1996 (first entry)
XX
XX Human thyroid transcription factor-1 gene.

XX Thyroid transcription factor; TTF-1; human adenocarcinoma cell line;
KW H441; rat; mouse; pulmonary adenocarcinoma; H820; small cell carcinoma;
KW H345; tracheal-bronchial epithelial cell lines; respiratory epithelium;
KW fecal lung; gestation; pro-SP-C; respiratory epithelial cell;
KW nonciliated bronchiolar cell; immature lung; alveolar; Type II;
KW epithelial cell; nonciliated; Type I; adult; ss.

XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH prim_transcript 143..3266
FT /*tag= a
FT exon 339..711
FT /*tag= b
FT intron 712..1674
FT /*tag= c
FT exon 1675..2417
FT /*tag= d
FT polyA_signal 3260..3266
FT /*tag= e

XX WO9531729-A1.
XX
XX 23-NOV-1995.
XX
XX 17-MAY-1995; 95WO-US006244.
XX
XX 18-MAY-1994; 94US-00245356.
XX
XX (CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.
XX
XX Bohinski RJ, Whitsett JA;
XX
XX WPI; 1996-011078/01.
XX
XX P-PSDB; AAR83015.

Oligo- and polynucleotide(s) that bind to lung cell nuclear proteins -
useful for cancer diagnosis and therapy.

Claim 50; Fig 39; 157pp; English.

XX This sequence encodes the thyroid transcription factor, TTF-1. This
sequence was isolated from the human adenocarcinoma cell line H441. The
TTF-1 locus is contained within a 4.6 kb BamHI fragment and consists of
two exons and one intron. The predicted amino acid sequence of human TTF-

PT New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
XX Claim 1; SEQ ID NO 7780; 41pp; English.
XX
XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein aduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
XX SQ Sequence 473 BP; 181 A; 77 C; 105 G; 110 T; 0 U; 0 Other;
Query Match 76.7%; Score 18.4; DB 13; Length 473;
Best Local Similarity 95.0%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 TAGTTAGCCTTTGGCTT 24
Db 308 TTGTTAGCCTTTGGCTT 289
RESULT 14
ADK16180/c
ID ADK16180 standard; DNA; 621 BP.
XX
XX AC ADK16180;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Nanoarchaeum equitans cancer-associated (CA) gene #66.
XX cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia;
XX ds; gene.
XX OS Nanoarchaeum equitans.
XX
XX PN WO2003093434-A2.
XX
XX PD 13-NOV-2003.
XX
XX PF 01-MAY-2003; 2003WO-US013699.
XX
XX PR 01-MAY-2002; 2002US-0377447P.
XX
XX PA (DIVE-) DIVERSA CORP.
XX Stetter KO, Waters E, Kretz K, Podar M, Richardson T;
XX
XX PI

PI Noordewier M;
XX WPI; 2004-053041/05.
DR P-PSDB; ADK16181.
XX
XX PT New recombinant cancer-associated genes, such as KCNJ9, useful for
PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,
PT cervical, or skin cancers, lymphomas, or leukemia.
XX
XX PS Claim 5; SEQ ID NO 132; 251pp; English.
XX
XX CC The invention comprises then amino acid and coding sequences of cancer-
CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention
CC also comprises the Nanoarchaeum equitans genome. The DNA and protein
CC sequences of the invention are useful for diagnosing and treating cancer
CC (e.g. carcinoma, lymphoma, or leukaemia). The present DNA sequence
CC represents a Nanoarchaeum equitans CA gene of the invention.
XX
XX SQ Sequence 621 BP; 232 A; 81 C; 86 G; 222 T; 0 U; 0 Other;
Query Match 75.8%; Score 18.2; DB 12; Length 621;
Best Local Similarity 87.0%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCAGTACTTTAGCCTTTGGCTT 23
Db 193 CGAATAGTTAGCCTTTGGCTT 171
RESULT 15
ADK16049 0
WP Sequence split into 5 fragments LOCUS ADK16049 Accession Adk16049
Fragment Name Begin End
WP ADK16049 0 1 110000
WP ADK16049 1 100001 210000
WP ADK16049 2 200001 310000
WP ADK16049 3 300001 410000
WP ADK16049 4 400001 490885
ID ADK16049 standard; DNA; 490885 BP.
XX
XX AC ADK16049;
XX
XX DT 06-MAY-2004 (first entry)
XX
XX DE Nanoarchaeum equitans genome.
XX
XX KW cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia;
XX ds.
XX OS Nanoarchaeum equitans.
XX
XX PN WO2003093434-A2.
XX
XX PD 13-NOV-2003.
XX
XX PF 01-MAY-2003; 2003WO-US013699.
XX
XX PR 01-MAY-2002; 2002US-0377447P.
XX
XX PA (DIVE-) DIVERSA CORP.
XX Stetter KO, Waters E, Kretz K, Podar M, Richardson T;
XX Noordewier M;
XX WPI; 2004-053041/05.
XX
XX PT New recombinant cancer-associated genes, such as KCNJ9, useful for
PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,
PT cervical, or skin cancers, lymphomas, or leukemia.
XX
XX PS Claim 5; SEQ ID NO 1; 251pp; English.
XX
XX CC The invention comprises then amino acid and coding sequences of cancer-

Search completed: July 20, 2005, 17:31:40
Job time : 225.652 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 16:53:52 ; Search time 75.6522 Seconds
(without alignments)
519.095 Million cell updates/sec

Title: US-10-041-030-8
Perfect score: 24
Sequence: 1 ccagtagtttagccttggctt 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfileseq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19.2	80.0	3293	2	US-08-442-809A-75
C 2	19.2	80.0	6250	4	US-09-949-016-13222
C 3	17.6	73.3	601	4	US-09-949-016-128695
C 4	17.6	73.3	235064	4	US-09-949-016-15390
C 5	17.4	72.5	205163	4	US-09-949-016-17009
C 6	17.2	71.7	601	4	US-09-949-016-158916
C 7	17.2	71.7	601	4	US-09-949-016-158917
C 8	17.2	71.7	601	4	US-09-949-016-158918
C 9	17.2	71.7	62354	4	US-09-949-016-16188
C 10	16.8	70.0	601	4	US-09-949-016-202241
C 11	16.8	70.0	790	1	US-08-306-691B-47
C 12	16.8	70.0	790	5	PCT-US93-06251-93
C 13	16.8	70.0	51161	4	US-09-949-016-17416
C 14	16.8	70.0	85869	4	US-09-949-016-12017
C 15	16.8	70.0	85878	4	US-09-949-016-16321
C 16	16.8	70.0	102304	4	US-09-949-016-12589
C 17	16.8	70.0	147382	4	US-09-949-016-14624
C 18	16.6	69.2	487	4	US-09-621-976-10207
C 19	16.6	69.2	1011	4	US-09-543-681A-1158
C 20	16.6	69.2	2004	4	US-09-252-991A-10609
C 21	16.6	69.2	2691	4	US-09-252-991A-10245
C 22	16.6	69.2	5976	3	US-08-621-944A-2
C 23	16.6	69.2	5976	3	US-08-945-567D-2
C 24	16.6	69.2	6426	4	US-09-540-236-1539
C 25	16.6	69.2	6973	1	US-08-478-370-1
C 26	16.6	69.2	6973	3	US-08-483-855C-1
C 27	16.6	69.2	6973	3	US-08-621-944A-1

C 28	16.6	69.2	6973	3	US-08-945-567D-1
C 29	16.6	69.2	6975	3	US-08-431-718C-1
C 30	16.6	69.2	9542	3	US-08-968-685A-9
C 31	16.6	69.2	62909	4	US-09-596-002-32
C 32	16.6	69.2	80269	4	US-09-949-016-15681
C 33	16.4	68.3	1680	4	US-09-328-352-1798
C 34	16.4	68.3	152486	4	US-09-949-016-12869
C 35	16.4	68.3	194933	4	US-09-949-016-14172
C 36	16.4	68.3	392000	4	US-10-027-983-11
C 37	16.2	67.5	223	4	US-09-549-848B-69
C 38	16.2	67.5	660	4	US-09-489-039A-4207
C 39	16.2	67.5	921	4	US-09-328-352-3953
C 40	16.2	67.5	1030	4	US-09-244-805-4
C 41	16.2	67.5	3630	3	US-09-221-017B-943
C 42	16.2	67.5	35100	2	US-08-770-379-18
C 43	16.2	67.5	35100	3	US-08-757-669A-18
C 44	16.2	67.5	35100	3	US-09-230-371A-18
C 45	16.2	67.5	44653	4	US-09-949-016-11944

ALIGNMENTS

RESULT 1
US-08-442-809A-75/c
; Sequence 75, Application US/08442809A
; Patent No. 5976873
; GENERAL INFORMATION:
; APPLICANT: Bohinski, Robert J.,
; APPLICANT: Whitsett, Jeffrey A.
; TITLE OF INVENTION: Nucleic Acid Sequences
; TITLE OF INVENTION: Controlling Lung Cell -
; TITLE OF INVENTION: Specific Gene Expression
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
; ADDRESS: Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,809A
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/245,356
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3293 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; FEATURE:
; NAME/KEY: human TTF-1 gene
; US-08-442-809A-75

Query Match 80.0%; Score 19.2; DB 2; Length 3293;

[illegible]

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158916
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-158916

Query Match          71.7%; Score 17.2; DB 4; Length 601;
Best Local Similarity 86.4%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AGTAGTTAGCCTTTGTGGCTT 24
Db 49 AGTAGTGTAGCCTTTTGGCGT 28

RESULT 7
US-09-949-016-158917/c
; Sequence 158917, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158917
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-158917

Query Match          71.7%; Score 17.2; DB 4; Length 601;
Best Local Similarity 86.4%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AGTAGTTAGCCTTTGTGGCTT 24
Db 329 AGTAGTGTAGCCTTTTGGCGT 308

RESULT 8
US-09-949-016-158918/c
; Sequence 158918, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158918
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-158918

Query Match          71.7%; Score 17.2; DB 4; Length 601;
Best Local Similarity 86.4%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AGTAGTTAGCCTTTGTGGCTT 24
Db 22904 AGTAGTGTAGCCTTTTGGCGT 22925

RESULT 9
US-09-949-016-16188
; Sequence 16188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16188
; LENGTH: 62354
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16188

Query Match          71.7%; Score 17.2; DB 4; Length 62354;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AGTAGTTAGCCTTTGTGGCTT 24
Db 22904 AGTAGTGTAGCCTTTTGGCGT 22925

RESULT 10
US-09-949-016-202241/c
; Sequence 202241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16188
; LENGTH: 62354
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202241/c

Query Match          71.7%; Score 17.2; DB 4; Length 62354;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AGTAGTTAGCCTTTGTGGCTT 24
Db 22904 AGTAGTGTAGCCTTTTGGCGT 22925
```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202241
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202241

```

```
Query Match          70.0%; Score 16.8; DB 4; Length 601;
Best Local Similarity 90.0%; Pred. No. 1e+02;
Matches 18: Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

RESULT 11
US-08-306-691B-47/c
; Sequence 47, Application US/08306691B
; Patent No. 5734039
;
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
;

```

Query Match	70.0%;	Score 16.8;	DB 1;	Length 790;
Best Local Similarity	90.0%;	Pred. No. 1.1e+02;		

	Matches	18;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	5	TAGTTTAGCCTTTGTGGCTT	24							
Dh	73	TACCTTACAGCTTTGTGGCTT	54							

```

RESULT 12
PCT-US93-06251-93/c
; Sequence 93, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 base pairs
; Type: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-93

```

Query Match	70.0%	Score 16.8;	DB 5;	Length 790;
Best Local Similarity	90.0%	Pred. No. 1.1e+02;		
Mismatches	18:	Conservative	0:	Mismatches 2:
				Indels 0:
				Gaps 0:

RESULT 13
US-09-949-016-17416/c
; Sequence 17416, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17416
; LENGTH: 51161
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17416

Query Match 70.0%; Score 16.8; DB 4; Length 51161;
Best Local Similarity 90.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CAGTAGTTTACCTTTGTGG 21
||| ||||| ||||| |||||
Db 32315 CAGGAGTTTAGACTTTGTGG 32296

RESULT 14

US-09-949-016-12017/c
; Sequence 12017, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12017
; LENGTH: 85869
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12017

Query Match 70.0%; Score 16.8; DB 4; Length 85869;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CAGTAGTTTACCTTTGTGG 21
||| ||||| ||||| |||||
Db 17435 CAATAGTTTAGGCTTTGTGG 17416

RESULT 15

US-09-949-016-16321/c
; Sequence 16321, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16321
; LENGTH: 85878

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16321

Query Match 70.0%; Score 16.8; DB 4; Length 85878;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CAGTAGTTTACCTTTGTGG 21
||| ||||| ||||| |||||
Db 17435 CAATAGTTTAGGCTTTGTGG 17416

Search completed: July 20, 2005, 18:58:58
Job time : 78.6522 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 18:01:31 ; Search time 337.043 Seconds
(without alignments)
451.758 Million cell updates/sec

Title: US-10-041-030-8
Perfect score: 24
Sequence: 1 ccagtagtttagccttggctt 24

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7173243 seqs, 3172129809 residues

Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	100.0	24	13	US-10-041-030-8
2	24	100.0	864	15	Sequence 8, Appli
3	24	100.0	3614	10	Sequence 1323, Ap
4	24	100.0	3615	10	Sequence 7043, Ap
5	24	100.0	3752	10	Sequence 7042, Ap
6	24	100.0	5579	14	Sequence 7041, Ap
7	24	100.0	5597	17	Sequence 135, App
					Sequence 287, App

ALIGNMENTS

RESULT 1
US-10-041-030-8
Sequence 8, Application US/10041030
Publication No. US20020150934A1
GENERAL INFORMATION:
APPLICANT: Powers, Scott
APPLICANT: Mu, David
APPLICANT: Xiang, Phil
APPLICANT: Feng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
FILE REFERENCE: 018781-006810US
CURRENT APPLICATION NUMBER: US/10/041,030
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: US 60/259,502
PRIOR FILING DATE: 2001-01-02
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide
OTHER INFORMATION: probe N63226R to pellino 2' untranslated region
US-10-041-030-8

Query Match 100.0%; Score 24; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.15;

Sequence 3, Appli
Sequence 286, App
Sequence 25, Appli
Sequence 75, Appli
Sequence 99303, A
Sequence 546, App
Sequence 254246,
Sequence 254247,
Sequence 254248,
Sequence 254249,
Sequence 254250,
Sequence 254251,
Sequence 254247,
Sequence 254248,
Sequence 254249,
Sequence 254250,
Sequence 254251,
Sequence 73, Appli
Sequence 19288, A
Sequence 60, Appli
Sequence 7, Appli
Sequence 6659, Ap
Sequence 45964, A
Sequence 45964, A
Sequence 58345, A
Sequence 58345, A
Sequence 58345, A
Sequence 1506, Ap
Sequence 1475, Ap
Sequence 1475, Ap
Sequence 10, Appli
Sequence 155, App
Sequence 1350, Ap
Sequence 119, App

Thu Jul 21 09:13:23 2005

us-10-041-030-8.rnpb

```

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTACGCTTTGTGGCTT 24
    |||||
Db 1 CCAGTAGTTTACGCTTTGTGGCTT 24
    |||||

RESULT 2
US-10-106-698-1323
; Sequence 1323, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1323
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (815)..(815)
; OTHER INFORMATION: n equals a.t,g, or c
; NAME/KEY: misc feature
; LOCATION: (848)..(848)
; OTHER INFORMATION: n equals a.t,g, or c
; NAME/KEY: misc feature
; LOCATION: (862)..(862)
; OTHER INFORMATION: n equals a.t,g, or c
US-10-106-698-1323

Query Match 100.0%; Score 24; DB 15; Length 864;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTACGCTTTGTGGCTT 24
    |||||
Db 260 CCAGTAGTTTACGCTTTGTGGCTT 283
    |||||

RESULT 3
US-09-764-891-7043
; Sequence 7043, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7043
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7043

Query Match 100.0%; Score 24; DB 10; Length 3614;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTACGCTTTGTGGCTT 24
    |||||
Db 3012 CCAGTAGTTTACGCTTTGTGGCTT 3035
    |||||

RESULT 4
US-09-764-891-7042
; Sequence 7042, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7042
; LENGTH: 3615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7042

Query Match 100.0%; Score 24; DB 10; Length 3615;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTACGCTTTGTGGCTT 24
    |||||
Db 3013 CCAGTAGTTTACGCTTTGTGGCTT 3036
    |||||

RESULT 5
US-09-764-891-7041
; Sequence 7041, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7041
; LENGTH: 3752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7041

Query Match 100.0%; Score 24; DB 10; Length 3752;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTACGCTTTGTGGCTT 24
    |||||
Db 3013 CCAGTAGTTTACGCTTTGTGGCTT 3036
    |||||

RESULT 6
US-10-197-666A-135
; Sequence 135, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAHI KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Elki phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18

```

; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 135
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (177)..(1436)
US-10-197-666A-135

Query Match 100.0%; Score 24; DB 14; Length 5579;
Best Local Similarity 100.0%; Pred. No. 0.31; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
|||||
Db 5062 CCAGTAGTTAGCCTTTGTGGCTT 5085

RESULT 7
US-10-085-117-287
; Sequence 287, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 5597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-287

Query Match 100.0%; Score 24; DB 17; Length 5597;
Best Local Similarity 100.0%; Pred. No. 0.31; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
|||||
Db 5062 CCAGTAGTTAGCCTTTGTGGCTT 5085

RESULT 8
US-10-041-030-3
; Sequence 3, Application US/10041030
; Publication No. US20020150934A1
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; FILE REFERENCE: Pellino Polypeptides and Polynucleotides
; FILE REFERENCE: 018781-006810US

; CURRENT APPLICATION NUMBER: US/10/041,030
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1763)
; OTHER INFORMATION: human pellino 2
US-10-041-030-3

Query Match 100.0%; Score 24; DB 13; Length 5921;
Best Local Similarity 100.0%; Pred. No. 0.31; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
|||||
Db 5386 CCAGTAGTTAGCCTTTGTGGCTT 5409

RESULT 9
US-10-085-117-286
; Sequence 286, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 202802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-286

Query Match 100.0%; Score 24; DB 17; Length 202802;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
|||||
Db 192286 CCAGTAGTTAGCCTTTGTGGCTT 192309

RESULT 10
US-10-236-031B-25/c
; Sequence 25, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullana, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102

Thu Jul 21 09:13:23 2005

us-10-041-030-8.rnpb

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 2202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-25

Query Match 80.0%; Score 19.2; DB 17; Length 2202;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
DB 1451 CCAGCAGTTTGGCCTTTGTGGTTT 1428

RESULT 11
US-09-320-337-75/c
; Sequence 75, Application US/09320337
; Patent No. US20010016352A1
; GENERAL INFORMATION:
; APPLICANT: Bohinski, Robert J.,
; APPLICANT: Whitsett, Jeffrey A.
; TITLE OF INVENTION: Nucleic Acid Sequences Controlling
; TITLE OF INVENTION: Lung Cell - Specific Gene Expression
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM P160
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/320,337
; FILING DATE: 26-MAY-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,809
; FILING DATE: 17-MAY-1995
; APPLICATION NUMBER: 08/245,356
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-447
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3293 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: oligonucleotide
; FEATURE:
; NAME/KEY: human TTF-1 gene
US-09-320-337-75

Query Match 80.0%; Score 19.2; DB 9; Length 3293;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
DB 2579 CCAGCAGTTTGGCCTTTGTGGTTT 2556

RESULT 12
US-10-424-599-99303/c
; Sequence 99303, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99303
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(364)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_60685C.1
US-10-424-599-99303

Query Match 75.8%; Score 18.2; DB 18; Length 364;
Best Local Similarity 87.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCAGTAGTTAGCCTTTGTGGCTT 23
DB 244 CCAGTAGTTGTGCTTGTGGCTT 222

RESULT 13
US-10-723-860-546
; Sequence 546, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 546
; LENGTH: 143239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-546

Query Match 75.8%; Score 18.2; DB 20; Length 143239;
Best Local Similarity 87.0%; Pred. No. 2.8e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CAGTAGTTAGCCTTTGTGGCTT 24
DB 31734 CAGTAATTTTGCCTTTGTGGCTT 31756

RESULT 14
US-10-027-632-254246
; Sequence 254246, Application US/10027632
; Publication No. US20020198371A1

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 254246
LENGTH: 1931
TYPE: DNA
ORGANISM: Human
US-10-027-632-254246

Query Match 74.2%; Score 17.8; DB 13; Length 1931;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCAGTAGTTAGCCTTTGTGG 21
Db 1461 CCAATAGTTAGCCTTTGTGG 1481

RESULT 15

US-10-027-632-254247
Sequence 254247, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 254247
LENGTH: 1931
TYPE: DNA
ORGANISM: Human
US-10-027-632-254247

Query Match 74.2%; Score 17.8; DB 13; Length 1931;
Best Local Similarity 90.5%; Pred. No. 2.4e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CCAGTAGTTAGCCTTTGTGG 21
Db 1461 CCAATAGTTAGCCTTTGTGG 1481
Search completed: July 20, 2005, 20:42:08
Job time : 340.043 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 11:59:56 ; Search time 929.739 Seconds
(without alignments)
1250.808 Million cell updates/sec

Title: US-10-041-030-8
Perfect score: 24
Sequence: 1 ccagtagtttagcctttgtgctt 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

Database : GenEmbl: *

```

1: gb_ba:
2: gb_htg
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts
12: gb_sy:
13: gb_un:
14: gb_v:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No./	Query	Score			Match	Length	DB	ID	Description
✓ 1		24	24	100.0	24	6	AX686410	Sequence	AX686410 Sequence
2		24	100.0	1890	9	AK025208	AK025208	Homo sapi	AK025208 Homo sapi
3		24	100.0	5579	6	BD190328	BD190328	Elk1 phos	BD190328 Elk1 phos
4		24	100.0	5597	9	AF302502	AF302502	Homo sapi	AF302502 Homo sapi
5		24	100.0	5921	6	AX686405	AX686405	Sequence	AX686405 Sequence
C 6		24	100.0	86794	2	AC016103	AC016103	Homo sapi	AC016103 Homo sapi
C 7		24	100.0	164550	9	CNS01RHV	CNS01RHV	Human chr	AL162633 Human chr
8		24	100.0	175988	9	CNS05TC7	CNS05TC7	Human chr	AL355073 Human chr
C 9		24	100.0	198278	2	AC010097	AC010097	Homo sapi	AC010097 Homo sapi
C 10		24	100.0	233753	2	AC137805	AC137805	Homo sapi	AC137805 Homo sapi
11		19.4	80.8	437	5	AB101392	AB101392	Haplochrom	AB101392 Haplochrom
12		19.4	80.8	236880	2	AC140762	AC140762	Rattus no	AC140762 Rattus no
13		19.4	80.8	247187	2	AC094928	AC094928	Rattus no	AC094928 Rattus no
C 14		19.2	80.0	165	9	AY083588	AY083588	Macaca mu	AY083588 Macaca mu
C 15		19.2	80.0	204	11	G29853	G29853	human STS S	G29853 human STS S
C 16		19.2	80.0	882	11	G26717	G26717	human STS S	G26717 human STS S
C 17		19.2	80.0	2131	9	BC006221	BC006221	Homo sapi	BC006221 Homo sapi
C 18		19.2	80.0	2202	9	HSU43203	HSU43203	Human thyr	U43203 Human thyr
C 19		19.2	80.0	2336	9	HSU33749	HSU33749	Human thyr	U33749 Human thyr

ALIGNMENTS

```

RESULT 1
AX686410 LOCUS linear PAT 29-MAR-2003
DEFINITION 24 bp DNA
ACCESSION Sequence 8 from Patent WO02059611.
AX686410
AX686410.1 GI:29372148
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
1
Powers,S., Mu,D., Xiang,P. and Peng,Y.
AUTHORS Diagnosis and treatment of cancer using mammalian pellino
TITLE polypeptides and polynucleotides
JOURNAL Patent: WO 02059611-A 8 01-AUG-2002;
Tularik Inc. (US)
FEATURES
source Location/Qualifiers
1..24
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="oligonucleotide probe N63226QR to pellino 2 3'
untranslated region"
ORIGIN
Query Match 100.0%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTAGCCTTTGGGCTT 24
|||||
DB 1 CCAGTAGTTTAGCCTTTGGGCTT 24

RESULT 2
AK025208 LOCUS linear PRI 13-SEP-2003
DEFINITION Homo sapiens cDNA: FLJ21555 fis, clone COL06351.
ACCESSION AK025208
AK025208.1 GI:10437673
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished

REFERENCE 2
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: fldcna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
source Location/Qualifiers
1. .1890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="COL06351"
/tissue_type="colon"
/clone_lib="COL"
/notes="cloning vector pME18SPL3"
238. .633
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15084.1"
/db_xref="GI:10437674"
/translation="MSCHFFKGLAPLPHVYTGNCRPIISCLGLTLPFASSPPEVKVP VMSYRNIFQLFMSFTKKKIQSGWSTLSIFLVRNLLIIGAAGSCMLQLSTPLFLG SAFSKGYSPPCRVLRFLDFLPTVAQKN"

CDS
238. .633
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15084.1"
/db_xref="GI:10437674"
/translation="MSCHFFKGLAPLPHVYTGNCRPIISCLGLTLPFASSPPEVKVP VMSYRNIFQLFMSFTKKKIQSGWSTLSIFLVRNLLIIGAAGSCMLQLSTPLFLG SAFSKGYSPPCRVLRFLDFLPTVAQKN"

ORIGIN
Query Match 100.0%; Score 24; DB 9; Length 1890;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24
|||||
Db 1303 CCAGTAGTTAGCCTTTGGCTT 1326

RESULT 3
BD190328
LOCUS BD190328 5579 bp DNA linear PAT 17-JUL-2003
DEFINITION Elkl phosphorylation related gene.
ACCESSION BD190328
VERSION BD190328.1 GI:33000067
KEYWORDS WO 03008589-A/68.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Matsuzaki, O., Matsuda, A., Nagano, Y. and Suzuki, N.
TITLE Elkl phosphorylation related gene
JOURNAL Patent: WO 03008589-A 68 30-JAN-2003;
ASAHI KASEI CORP, OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI
COMMENT OS Homo sapiens (human)
PN WO 03008589-A/68
PD 30-JAN-2003

PF 15-JUL-2002 WO 2002JP007174
PR 18-JUL-2001 JP 01P 218204, 31-AUG-2001 JP 01P 263450 PR
21-JAN-2002 JP 02P 012176
PI OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
C12N15/54, C12N19/12, G01N33/50, G01N33/15, C07K16/40, A61K39/385, PC
A61K31/711
CC Elkl phosphorylation related gene
FH Key Location/Qualifiers
FT CDS (177). .(1436).
FEATURES
source Location/Qualifiers
1. .5579
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 24; DB 6; Length 5579;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24
|||||
Db 5062 CCAGTAGTTAGCCTTTGGCTT 5085

RESULT 4
AF302502
LOCUS AF302502 5597 bp mRNA linear PRI 05-JUL-2001
DEFINITION Homo sapiens pellino 2 (PELLI2) mRNA, complete cds.
ACCESSION AF302502
VERSION AF302502.1 GI:10242352
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLE Assignment of homologous genes, Pelli1/PELLI1 and Pelli2/PELLI2, for the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human chromosomes 2p13.3 and 14q21, respectively, by physical and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
MEDLINE 21203570
PUBMED 11306823
REFERENCE 2 (bases 1 to 5597)
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular Pathology, University of Bielefeld, Universitaetsstrasse 25, Bielefeld, NRW 33615, Germany
FEATURES
source Location/Qualifiers
1. .5597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .5597
/gene="PELLI2"
177. .1439
/gene="PELLI2"
/note="similar to Drosophila melanogaster Pellino"
/codon_start=1
/product="pellino 2"
/protein_id="AAG15390.1"
/db_xref="GI:10242353"
/translation="MFSPQBEHCAPNKEPVKYGLVVLGYNGALPNGDRGRKRRPFA LYKRPKANGVKPSTVHVITSPQASKAISCKQHSISYTLNRNQTVVVYTHDKDMDM QVGRSTESPIDFVVTDTTSGQNTDEAOITOSTSRFACRIVCDRNEPVYTFARIFAAG DSSKNIFTEGKAALKWKNPDGMDGLTTNGVLVMPHPRGGFTESQGVNREISVCCDGV TUREARQQRKLVSESTNVLDGSLDLCATLLWPTADGLFHTPTQKHLEALROE INARAPQCPLNTLAFPSINRKEVSEKQWYSLSCGHVGHVHNGHRSRSDTEANERE CPMCRVTGPPVPLWLGCSAGFYVDAGPPHTAFTPCGHVCSKSAKYSQIPLPHGTHA FHACPFCAQTQLVGEQNCIKLIFQGPID"


```
* 16029 16893: contig of 865 bp in length
* 16894 16931: gap of 100 bp
* 16984 17861: contig of 868 bp in length
* 17862 17961: gap of 100 bp
* 17962 18856: contig of 895 bp in length
* 18857 18956: gap of 100 bp
* 18957 19822: contig of 866 bp in length
* 19823 19922: gap of 100 bp
* 19923 20840: contig of 918 bp in length
* 20841 20940: gap of 100 bp
* 20941 21855: contig of 915 bp in length
* 21856 21955: gap of 100 bp
* 21956 22811: contig of 856 bp in length
* 22812 22911: gap of 100 bp
* 22912 23825: contig of 914 bp in length
* 23826 23925: gap of 100 bp
* 23926 24838: contig of 912 bp in length
* 24839 24937: gap of 100 bp
* 24938 25848: contig of 911 bp in length
* 25849 26846: contig of 898 bp in length
* 26847 26946: gap of 100 bp
* 26947 27844: contig of 898 bp in length
* 27845 27944: gap of 100 bp
* 27945 28802: contig of 858 bp in length
* 28803 28902: gap of 100 bp
* 28903 29817: contig of 915 bp in length
* 29818 29917: gap of 100 bp
* 29918 30805: contig of 888 bp in length
* 30806 31808: contig of 903 bp in length
* 31809 32805: contig of 897 bp in length
* 32806 32905: gap of 100 bp
* 32906 33766: contig of 861 bp in length
* 33767 33866: gap of 100 bp
* 33867 34762: contig of 896 bp in length
* 34763 34862: gap of 100 bp
* 34863 35762: contig of 900 bp in length
* 35763 35862: gap of 100 bp
* 35863 36770: contig of 908 bp in length
* 36771 36870: gap of 100 bp
* 36871 37832: contig of 962 bp in length
* 37833 37932: gap of 100 bp
* 37933 38814: contig of 882 bp in length
* 38815 38914: gap of 100 bp
* 38915 39783: contig of 869 bp in length
* 39784 39883: gap of 100 bp
* 39884 40823: contig of 940 bp in length
* 40824 40923: gap of 100 bp
* 40924 41823: contig of 900 bp in length
* 41824 41923: gap of 100 bp
* 41924 42808: contig of 885 bp in length
* 42809 42908: gap of 100 bp
* 42909 43822: contig of 914 bp in length
* 43823 43922: gap of 100 bp
* 43923 44809: contig of 887 bp in length
* 44810 44909: gap of 100 bp
* 44910 45779: contig of 870 bp in length
* 45780 45879: gap of 100 bp
* 45880 46772: contig of 893 bp in length
* 46773 46872: gap of 100 bp
* 46873 47758: contig of 885 bp in length
* 47759 47858: gap of 100 bp
* 47859 48778: contig of 920 bp in length
* 48779 48878: gap of 100 bp
* 48879 49740: contig of 862 bp in length
* 49741 49840: gap of 100 bp
* 49841 50694: contig of 854 bp in length
* 50695 50794: gap of 100 bp
* 50795 51699: contig of 905 bp in length
* 51700 51799: gap of 100 bp
* 51800 52685: contig of 886 bp in length
```

```
* 52686 52785: gap of 100 bp
* 52786 53705: contig of 920 bp in length
* 53706 53805: gap of 100 bp
* 53806 54718: contig of 913 bp in length
* 54719 54818: gap of 100 bp
* 54819 55719: contig of 901 bp in length
* 55720 55819: gap of 100 bp
* 55820 56733: contig of 914 bp in length
* 56734 56833: gap of 100 bp
* 56834 57742: contig of 909 bp in length
* 57743 57842: gap of 100 bp
* 57843 58753: contig of 911 bp in length
* 58754 58853: gap of 100 bp
* 58854 59748: contig of 895 bp in length
* 59749 59848: gap of 100 bp
* 59849 60720: contig of 872 bp in length
* 60721 60820: gap of 100 bp
* 60821 61697: contig of 877 bp in length
* 61698 61797: gap of 100 bp
* 61798 62699: contig of 902 bp in length
* 62700 62799: gap of 100 bp
* 62800 63701: contig of 902 bp in length
* 63702 64667: contig of 866 bp in length
* 64668 64767: gap of 100 bp
* 64768 65673: contig of 906 bp in length
* 65674 65773: gap of 100 bp
* 65774 66698: contig of 925 bp in length
* 66699 66798: gap of 100 bp
* 66799 67694: contig of 896 bp in length
* 67695 67794: gap of 100 bp
* 67795 68697: contig of 903 bp in length
* 68698 68797: gap of 100 bp
* 68798 69697: contig of 900 bp in length
* 69698 69797: gap of 100 bp
* 69798 70737: contig of 940 bp in length
* 70738 70837: gap of 100 bp
* 70838 71729: contig of 892 bp in length
* 71730 71829: gap of 100 bp
* 71830 72713: contig of 884 bp in length
```

Query Match 100.0%; Score 24; DB 2; Length 86794;

Best Local Similarity 100.0%; Pred. No. 0.54;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGTGCTT 24

Db 13221 CCAGTAGTTAGCCTTTGTGCTT 13198

RESULT 7

CNS01RHY/c

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Reference

Authors

Title

Journal

CNS01RHY Human chromosome 14 DNA sequence BAC R-21008 of library RPCI-11 from chromosome 14 of Homo sapiens (Human), complete sequence.

AL162633

AL162633.3 GI:11545121

HTG.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 164550)

Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,

Brottier, P., Cattoico, L., Barbe, V., Pelletier, E., Attiguenave, F.,

Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,

Guay, G., Saurin, W. and Weissenbach, J.

Sequencing of the human chromosome 14

Unpublished

2 (bases 1 to 164550)

Genoscope.

Direct Submission

Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 On Dec 4, 2000 this sequence version replaced gi:17799785.
 ----- Genome Center
 Center: Genoscope / Centre National de Sequencage
 Center code: GS
 Web site: http://www.genoscope.cns.fr/
 Contact: SeqRef@genoscope.cns.fr

COMMENT

The following BAC sequence is oriented from the T7 to the SP6 end.
 Upstream BAC (overlapping the T7 end) : R-930011
 Downstream BAC (overlapping the SP6 end) : C-2184N20 (AC=AL359234)
 ----- Summary Statistics
 Assembly program: Phrap; version 2.0
 Quality coverage: 16.67x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases

0 : 1 - 9 : 1

10 - 19 : 8

20 - 29 : 19

30 - 39 : 125

40 - 49 : 742

50 - 59 : 1376

60 - 69 : 1816

70 - 79 : 4694

80 - 89 : 21302

90 - 99 : 134467

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

source

1. 164550

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="14"

/clone_lib="RPCI-11"

142293..142420

/notes="matching EMBL:H53462"

RHdb:RH53794

dbSTS:STS6714

Identified using the e-PCR software (G. Schuler)"

STS

ORIGIN

Query Match 100.0%; Score 24; DB 9; Length 164550;
 Best Local Similarity 100.0%; Pred. No. 0.5;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24

Db 112410 CCAGTAGTTAGCCTTTGGCTT 112387

RESULT 8

CNS05TC7

LOCUS

DEFINITION

Human chromosome 14 DNA sequence BAC R-930011 of library RPCI-11

from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION AL355073

VERSION 1

KEYWORDS HTG.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 175988)

AUTHORS Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,

Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,

Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,

Gyapay, G., Saurin, W. and Weissensbach, J.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Sequencing of the human chromosome 14

Unpublished

2 (bases 1 to 175988)

Genoscope.

Direct Submission

Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

On Apr 30, 2001 this sequence version replaced gi:12697154.

----- Genome Center

Center: Genoscope / Centre National de Sequencage

Center code: GS

Web site: http://www.genoscope.cns.fr/

Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.

Upstream BAC (overlapping the T7 end) : C-2588E21

Downstream BAC (overlapping the SP6 end) : R-21008 (AC=AL162633)

----- Summary Statistics

Assembly program: Phrap; version 2.0

Quality coverage: 7.58x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases

0 : 1 - 9 : 1

10 - 19 : 8

20 - 29 : 19

30 - 39 : 125

40 - 49 : 742

50 - 59 : 1376

60 - 69 : 1816

70 - 79 : 4694

80 - 89 : 21302

90 - 99 : 101037

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

source

1. 175988

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="14"

/clone_lib="RPCI-11"

72485..72612

/note="matching EMBL:H53462"

RHdb:RH53794

dbSTS:STS6714

Identified using the e-PCR software (G. Schuler)"

ORIGIN

Query Match

Best Local Similarity

Matches 24;

Conservative 0;

Mismatches 0;

Indels 0;

Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24

Db 102488 CCAGTAGTTAGCCTTTGGCTT 102511

RESULT 9

AC010097/c

LOCUS

DEFINITION

Homo sapiens chromosome 14 clone RP11-378116, WORKING DRAFT

SEQUENCE, 18 unordered pieces.

ACCESSION AC010097

VERSION AC010097.6

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AC010097 198278 bp DNA linear HTG 07-JUL-2000

Homo sapiens chromosome 14 clone RP11-378116, WORKING DRAFT

SEQUENCE, 18 unordered pieces.

ACCESSION AC010097

VERSION AC010097.6

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 198278)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 198278)
Waterston,R.H.
Direct Submission
Submitted (11-SEP-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 7, 2000 this sequence version replaced gi:8568179.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.NH0378116
----- Summary Statistics -----
Sequencing vector: M13; 88%
Sequencing vector: plasmid; 12%
Chemistry: Dye-terminator Big Dye; 27% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 191558 bases at least Q40
Consensus quality: 193468 bases at least Q30
Consensus quality: 194761 bases at least Q20
Insert size: 195000; agarose-fp
Quality coverage: 5.76 in Q20 bases; agarose-fp
Quality coverage: 5.18 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1211: contig of 1211 bp in length
* 1212 1311: gap of unknown length
* 1312 2946: contig of 1635 bp in length
* 2947 3047 6236: gap of unknown length
* 6237 6336: gap of unknown length
* 6337 9363: contig of 3027 bp in length
* 9364 9463: gap of unknown length
* 9464 13996: contig of 4532 bp in length
* 13996 14096: gap of unknown length
* 14096 17948: contig of 3853 bp in length
* 17949 18048: gap of unknown length
* 18049 23265: contig of 5217 bp in length
* 23266 23365: gap of unknown length
* 23366 30448: contig of 7083 bp in length
* 30449 30549: gap of unknown length
* 30549 41508: contig of 10960 bp in length
* 41509 51502: contig of 9894 bp in length
* 51503 61316: gap of unknown length
* 61317 72947: contig of 11531 bp in length
* 72948 73047: gap of unknown length
* 73048 87328: contig of 14280 bp in length
* 87329 87428: gap of unknown length
* 87429 103771: contig of 16344 bp in length
* 103772 103871: gap of unknown length
* 103872 120942: contig of 17071 bp in length
* 120943 138875: contig of 17833 bp in length
* 138876 138975: gap of unknown length

* 138976 162621: contig of 23646 bp in length
* 162622 162721: gap of unknown length
* 162722 198278: contig of 35557 bp in length.
FEATURES
source
1.198278
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="RP11-378116"
ORIGIN
Query Match 100.0% Score 24; DB 2; Length 198278;
Best Local Similarity 100.0%; Pred.No. 0.49; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;
QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
|||||
Db 106197 CCAGTAGTTAGCCTTTGTGGCTT 106174
|||||
RESULT 10
AC137805/c
LOCUS
DEFINITION Homo sapiens chromosome 16 clone RP11-863C20, WORKING DRAFT
SEQUENCE, 2 unordered pieces.
ACCESSION AC137805
VERSION AC137805.1 GI:26006547
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 233753)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 233753)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1625964
Center clone name: RPCI-11_863C20

Summary Statistics
Consensus quality: 233513 bases at least Q40
Consensus quality: 233587 bases at least Q30
Consensus quality: 233607 bases at least Q20
Estimated insert size: 160000; agarose-fp estimation
Quality coverage: 18.97 in Q20 bases; agarose-fp estimation
Quality coverage: 12.99 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 23031: contig of 23031 bp in length
* 23032 23131: gap of unknown length
* 23132 233753: contig of 210622 bp in length.
FEATURES
source
1.233753
Location/Qualifiers
/organism="Homo sapiens"

ORIGIN

Query Match 100.0%; Score 24; DB 2; Length 233753;
 Best Local Similarity 100.0%; Pred. No. 0.48; 0; Gaps 0;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 1 CCAGTACTTTAGCCTTTGTGGCTT 24
 |||
 Db 131315 CCAGTACTTTAGCCTTTGTGGCTT 131292

HTG: HTGS PHASE1: HTGS DRAFT: HTGS_ENRICHED.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

RESULT 11

AB101392

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 236880)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angulano, D., Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Eacotto, M., Eugene, C., Evans, C. A., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Loulaeged, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelamen, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, C., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiczysk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished

2 (bases 1 to 236880)

Worley, K. C.

Direct Submission

Submitted (01-MAR-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 236880)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:28626613.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

COMMENT

/mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-863C20"
 /clone_lib="RPC1 human BAC library 11"

AB101392

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (28-JAN-2003) Yohey Terai, Tokyo Institute of Technology, Graduate School of Bioscience and Biotechnology; 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8501, Japan (E-mail: yterai@bio.titech.ac.jp, Tel: 81-45-924-5744, Fax: 81-45-924-5835)

repeat_region 184..437
 /notes="1322 locus for SINE sequence analysis"
 /rpt_family="AFC"
 /rpt_type="dispersed"

FEATURES

source

1..437
 /organism="Haplochromis nyererei"
 /mol_type="genomic DNA"
 /db_xref="taxon:122282"
 /tissue_type="piece of fin"

repeat_region 184..437
 /notes="1322 locus for SINE sequence analysis"
 /rpt_family="AFC"
 /rpt_type="dispersed"

ORIGIN

Query Match 80.8%; Score 19.4; DB 5; Length 437;
 Best Local Similarity 95.2%; Pred. No. 1.8e+02;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AGTAGTTTAGCCTTTGTGGCT 23
 |||
 Db 169 AGTTGTTTAGCCTTTGTGGCT 189

RESULT 12

AC140762

LOCUS

DEFINITION

ACCESSION

VERSION

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:28626613.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GK2B
Center clone name: CH230-70P5
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 207907 bases at least Q40
Consensus quality: 211702 bases at least Q30
Consensus quality: 214041 bases at least Q20
Estimated insert size: 216628; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 3 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
*
* 1 234297: contig of 234297 bp in length
* 234298 234397: gap of unknown length
* 234398 235412: contig of 1015 bp in length
* 235413 235512: gap of unknown length
* 235513 236880: contig of 1368 bp in length.
FEATURES             Location/Qualifiers
source               1..236880
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-70P5"
misc_feature         1..950
                    /note="wgs_contig"
ORIGIN
Query Match      80.8%; Score 19.4; DB 2; Length 236880;
Best Local Similarity 95.2%; Pred. No. 84;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CCAGTAGTTAGCCTTTGTGG 21
|||||
Db 11172 CCAGTAGTTAGCCTTTGTGG 11192

```

```

RESULT 13
AC094928
LOCUS
DEFINITION
AC094928 Rattus norvegicus clone CH230-6D14, WORKING DRAFT SEQUENCE, 4
UNORDERED
AC094928
VERSION
AC094928.7 GI:30466913
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE

AUTHORS

1 (bases 1 to 247187)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsebrook, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, W., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havliak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., L., Loulseg, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakeleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, I., Rojals, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Woodden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 247187)

Worley, K. C.

Direct Submission

Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 247187)

Rat Genome Sequencing Consortium

Direct Submission

Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 9, 2003 this sequence version replaced gi:23264792.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GBTD

Center clone name: CH230-6D14

----- Summary Statistics

Assembly program: Atlas;

Consensus quality: 228892 bases at least Q40

Consensus quality: 231505 bases at least Q30

Consensus quality: 233248 bases at least Q20

Estimated insert size: 228978; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

- * NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
- * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

- * 1 241575: contig of 241575 bp in length
- * 241576 241675: gap of unknown length
- * 241676 242742: contig of 1067 bp in length
- * 242743 242842: gap of unknown length
- * 242843 244551: contig of 1709 bp in length
- * 244552 244651: gap of unknown length
- * 244652 247187: contig of 2536 bp in length.

FEATURES

source
1. .247187
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-6D14"
1. .1736
/notes="wgs_end_extension
clone_end:Sp6"
complement(4508..5181)
/notes="clone_boundary
clone_end:Sp6
site:EcoRI
end_sequence:BH358097"
complement(232634..233337)
/note="clone_boundary
clone_end:T7
site:EcoRI
end_sequence:BH358095"
234739..235881
/note="wgs_end_extension
clone_end:T7"
239927..241575
/note="wgs_end_extension
clone_end:T7"

ORIGIN

Query Match 80.8%; Score 19.4; DB 2; Length 247187;
Best Local Similarity 95.2%; Pred. No. 84;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGTGG 21
|||||

Db 136174 CCAGTAGTTAGCCTTTGGTGG 136194
|||||

RESULT 14
AY083588/c

LOCUS 165 bp DNA linear PRI 16-APR-2002
Macaca mulatta thyroid transcription factor 1 (TTFP1) gene, 3' UTR.

DEFINITION AY083588

ACCESSION AY083588.1 GI:20159658

VERSION AY083588.1

KEYWORDS

SOURCE

ORGANISM

Macaca mulatta (rhesus monkey)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopithecoidea; Cercopithecoidea; Cercopithecoidea;

1 (bases 1 to 165)

Norgren,R.B. Jr., Zink,M.A., Jia,Y., Ojeda,S.R. and Spindel,E.R.

Construction of a targeted rhesus macaque microarray

Unpublished

2 (bases 1 to 165)

Norgren,R.B. Jr., Zink,M.A., Jia,Y., Ojeda,S.R. and Spindel,E.R.

Direct Submission

Submitted (11-MAR-2002) Molecular and Cellular Biology Core, Oregon

Regional Primate Research Center, 505 NW 185th Avenue, Beaverton,

OR 97006, USA

FEATURES

Location/Qualifiers

1. .165

/organism="Macaca mulatta"

/mol_type="genomic DNA"

/db_xref="taxon:9544"

<1..>165

/gene="TTFP1"

<1..>165

/gene="TTFP1"

/product="thyroid transcription factor 1"

<1..>165

/gene="TTFP1"

3' UTR

Query Match 80.0%; Score 19.2; DB 9; Length 165;

Best Local Similarity 87.5%; Pred. No. 2.6e+02;

Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24
|||||

Db 56 CCAGCAGTTTGGCTTTGGCTT 33
|||||

RESULT 15
G29853/c

LOCUS 204 bp DNA linear STS 05-OCT-1996
human STS SHGC-35528, sequence tagged site.

DEFINITION G29853

ACCESSION G29853

VERSION G29853.1 GI:1593404

KEYWORDS STS; STS sequence; primer; sequence tagged site.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 204)

Myers,R.M.

Unpublished (1996)

REFERENCE

AUTHORS

JOURNAL

COMMENT

Contact: Richard M. Myers

Stanford Human Genome Center (SHGC)

Stanford University School of Medicine

Department of Genetics, M-344, Stanford, CA 94305, USA

Tel: 4157259687

Fax: 4157259689

Email: myers@shgc.stanford.edu

Primer A: ACAGCTTCGACTTTCTTA

Primer B: CCAAGCTGTTTATGCCCT

STS size: 204

PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol: Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from U33749
-- Washington University/Merck EST sequence.

FEATURES

source
1..204
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="14"
STS
1..204
primer_bind
1..20
primer_bind
complement(185..204)

ORIGIN

Query Match 80.0%; Score 19.2; DB 11; Length 204;
Best Local Similarity 87.5%; Pred. No. 2.5e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCAGTAGTTTAGCCTTTGTGGCTT 24
Db 77 CCAGCAGTTTGGCCTTTGTGGTTT 54

Search completed: July 20, 2005, 18:01:27
Job time : 934.739 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 11:59:56 ; Search time 852.261 Seconds
(without alignments)
1250.808 Million cell updates/sec

Title: US-10-041-030-7

Perfect score: 22

Sequence: 1 gatgctgaagtcgtctcattgg 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_bcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sbs.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	6	AX686409 Sequence
2	22	100.0	60	6	CQ543281 Sequence
3	22	100.0	5579	6	BD190328 Elk1 phos
4	22	100.0	5579	6	AF302502 Homo sapi
5	22	100.0	5921	6	AX686405 Sequence
6	22	100.0	86794	2	AC016103 Homo sapi
7	22	100.0	164550	9	CNS01RHY
8	22	100.0	175988	2	CNS05TC7
9	22	100.0	198278	2	AC010097
10	22	100.0	233753	2	AC137805
11	20.4	92.7	1890	9	AK025208
12	17.8	80.9	539	8	HAN507474
13	17.8	80.9	549	8	HAN507473
14	17.8	80.9	550	8	HIN507471
15	17.8	80.9	568	8	HPA507469
16	17.8	80.9	688	8	AF173650
17	17.8	80.9	163246	2	AC120248
18	17.8	80.9	174021	9	AL358612
19	17.8	80.9	178004	2	AL590070 Homo sapi

c	20	17.8	80.9	220052	5	BX927308
	21	17.8	80.9	247754	2	AC106199
	22	17.8	80.9	249418	2	AC097209
	23	17.4	79.1	134787	2	AL356692
c	24	17.4	79.1	164529	2	AC079975
	25	17.4	79.1	172482	10	AC115060
	26	17.4	79.1	172753	9	AL359259
	27	17.4	79.1	176671	2	AC080149
c	28	17.4	79.1	203279	9	AL354826
	29	17.2	78.2	579	6	AR557605
c	30	17.2	78.2	867	1	AB014978
	31	17.2	78.2	1485	4	BTTP1G
c	32	17.2	78.2	2014	8	APU12757
	33	17.2	78.2	2182	6	AX427149
	34	17.2	78.2	3369	9	HSM805437
	35	17.2	78.2	3911	9	HSM806334
	36	17.2	78.2	3992	9	HSM806209
	37	17.2	78.2	4130	6	BD191477
	38	17.2	78.2	4218	9	HSM806037
	39	17.2	78.2	4458	9	HSM806297
	40	17.2	78.2	4799	6	CQ843490
	41	17.2	78.2	4799	9	AK126472
c	42	17.2	78.2	7167	1	AE011527
	43	17.2	78.2	9217	9	AB095939
	44	17.2	78.2	10867	1	AE001594
c	45	17.2	78.2	10871	1	AE002227

ALIGNMENTS

RESULT 1	AX686409	22 bp	DNA	linear	PAT 29-MAR-2003
LOCUS	Sequence 7 from Patent WO02059611.				
DEFINITION	AX686409				
ACCESSION	AX686409				
VERSION	AX686409.1	GI:29372147			
KEYWORDS	synthetic construct				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				

REFERENCE 1

AUTHORS Powers,S., Mu,D., Xiang,P. and Peng,Y.

TITLE Diagnosis and treatment of cancer using mammalian pellino

JOURNAL polypeptides and polynucleotides

PATENT: WO 02059611-A 7 01-AUG-2002;

Tularik Inc. (US)

Location/Qualifiers

1..22

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32830"

/note="oligonucleotide probe N632260F to pellino 2 3'

untranslated region"

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 22;

Best Local Similarity 100.0%; Pred.No. 2.2;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTGG 22

|||||

Db 1 GATGCTGAAGTCGTCCTCATTGG 22

|||||

RESULT 2

LOCUS CQ543281/c

DEFINITION Sequence 12916 from Patent WO0210449.

ACCESSION CQ543281

VERSION CQ543281.1

KEYWORDS Homo sapiens (human)

SOURCE

60 bp

DNA

linear

PAT 30-JAN-2004

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Shohan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
AUTHORS Oligonucleotide library for detecting rna transcripts and splice
TITLE variants that populate a transcriptome
JOURNAL Patent: WO 0210449-A 12916 07-FEB-2002;
CompuGen Inc. (US)
FEATURES
source
1..60
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 22; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATGCTGAAGTCGTCCTCATTTG 22
Db 60 GATGCTGAAGTCGTCCTCATTTG 39
RESULT 3
BD190328/c
LOCUS BD190328 5579 bp DNA linear PAT 17-JUL-2003
DEFINITION Elki phosphorylation related gene.
ACCESSION BD190328
VERSION BD190328.1 GI:33000067
KEYWORDS WO 03008589-A/68.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Matsuzaki,O., Matsuda,A., Nagano,Y. and Suzuki,N.
AUTHORS Elki phosphorylation related gene
TITLE Patent: WO 03008589-A 68 30-JAN-2003;
JOURNAL ASAHII KASEI CORP, OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI
COMMENT OS Homo sapiens (human)
PN WO 03008589-A/68
PD 30-JAN-2003
PF 15-JUL-2002 WO 2002JP007174
PR 18-JUL-2002 JP 01P 218204,31-AUG-2001 JP 01P 263450 PR
PI OSAMU MATSUZAKI,AKIO MATSUDA, YUKIKO NAGANO,NAOMI SUZUKI PC
C12N15/54,C12N9/12,G01N33/50,G01N33/15,C07K16/40,A61K39/385, PC
A61K31/711
CC Elki phosphorylation related gene
FH Key Location/Qualifiers
FT CDS (177)..(1436).
FEATURES
source
1..5579
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 22; DB 6; Length 5579;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATGCTGAAGTCGTCCTCATTTG 22
Db 5135 GATGCTGAAGTCGTCCTCATTTG 5114
RESULT 4
AF102502/c
LOCUS AF102502 5597 bp mRNA linear PRI 05-JUL-2001

```

```

DEFINITION Homo sapiens pellino 2 (PELLI2) mRNA, complete cds.
ACCESSION AF302502
VERSION AF302502.1 GI:10242352
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Resch,K., Jockusch,H. and Schmitt-John,T.
AUTHORS Assignment of homologous genes, Pelli1/PELLI1 and Pelli2/PELLI2, for
TITLE the Pelli adaptor protein pellino to mouse chromosomes 11 and 14
and human chromosomes 2p13.3 and 14q21, respectively, by physical
and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
MEDLINE 21203570
PUBMED 11306823
REFERENCE 2 (bases 1 to 5597)
AUTHORS Resch,K., Jockusch,H. and Schmitt-John,T.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular
Pathology, University of Bielefeld, Universitaetsstrasse 25,
Bielefeld, NRW 33615, Germany
FEATURES
source
1..5597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..5597
/gene="PELLI2"
177..1439
/gene="PELLI2"
/note="similar to Drosophila melanogaster Pellino"
/codon_start=1
/product="pellino 2"
/protein_id="AAG15390.1"
/db_xref="GI:10242353"
/translation="MFSQGEHCAPNKEPVKYGELVVLVNGALPNGDRGRKSRFA
LYKRPKANGVPSTVHVISTPOAKAISCKQHSISYTLNRQTQVVEYTHDKDITDMF
QVGRSTSPIDFVVTDTTIGSQNTDEAQTQSTISRACRIVCDRNEPYTARIFAAGF
DSSKNIFGEKAARKWNPDGHMDGLTTNGVLVMPHPRGFTBESQGVVREISVCGDVY
TLRETRSAQRQKLVESNTVLQDGLIDLGCATLLWRTADGLPHTPTQKHLEALRQE
INAAEPCPGVGLNTLAFPSINRKEVVEEKQKAVLSCGHVGHYHNGHRSSTEANERE
CPMCTVGPVPLWLGCEAGFYVDAGPTPCTGCHVCSKSAKISQIPLPHGTHA
FHAACPFCAATQLVGEQNCIKLIFQGPID"
ORIGIN
Query Match 100.0%; Score 22; DB 9; Length 5597;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATGCTGAAGTCGTCCTCATTTG 22
Db 5135 GATGCTGAAGTCGTCCTCATTTG 5114
RESULT 5
AX686405/c
LOCUS AX686405 5921 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 3 from Patent WO02059611.
ACCESSION AX686405
VERSION AX686405.1 GI:29372143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Powers,S., Mu,D., Xiang,P. and Peng,Y.
AUTHORS Diagnosis and treatment of cancer using mammalian pellino
TITLE polypeptides and polynucleotides
JOURNAL Patent: WO 02059611-A 3 01-AUG-2002;
Tularik Inc. (US)

```

```

FEATURES             source
1. .5921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
501..1763
/notes="unnamed protein product; human pellino 2"
/codon_start=1
/db_xref="GI:29372144"
/translation="MFSPGOEEHCAPNKPVKYGLVGLVNGALPNGDRGRKSRPA
LYKPKANGKVPSTVHVISPOASKALSCQGHISITLSRNQTVVVEYTHDKDTMF
QVGRSTSPDVFVDTISGSDTEAQITQSTISRFACRIVCDRNEPYTHARIFAAG
DSSKNIIFLGEAKAKNPDGMDGLTTNGVLVMPHPRGGFTESQGVYREISVCGDYV
TLRETRAQORGLKVESETNVLQDGLIDLCATLLWRTADGLPHTPTQKHIALROE
INAAPOCPVGLNTLAFPSINRKEVVEEKOPWALSCGHVGHVHNGHRSRDTANERE
CPMCRTVGPVPLWLGCEAGFYVDAGPTTHAFTPCGHVCEKSKAKYWSQILPHGTHA
FHAACPCATQLVGEQNCIKLIFQPID"

ORIGIN
      100.0%; Score 22; DB 6; Length 5921;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCATTGG 22
      |||||
Db 5459 GATGCTGAAGTCGTCATTGG 5438

RESULT 6
AC016103
LOCUS
DEFINITION
AC016103
VERSION
AC016103.2 GI:9134395
KEYWORDS
HTG; HTGS_PHASEO.
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 86794)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-26H17
Unpublished
2 (bases 1 to 86794)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6456211.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4637

```

Center clone name: 26_H_17

NOTE: This record contains 87 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

```

      1      878: contig of 878 bp in length
      879      978: gap of 100 bp
      979      1882: contig of 904 bp in length
      1883      1982: gap of 100 bp
      1983      2880: contig of 898 bp in length
      2881      2980: gap of 100 bp
      2981      3900: contig of 920 bp in length
      3901      4000: gap of 100 bp
      4001      4856: contig of 856 bp in length
      4857      4956: gap of 100 bp
      4957      5857: contig of 901 bp in length
      5858      5957: gap of 100 bp
      5959      6873: contig of 916 bp in length
      6874      6973: gap of 100 bp
      6974      7854: contig of 881 bp in length
      7855      7954: gap of 100 bp
      7955      8846: contig of 892 bp in length
      8847      8946: gap of 100 bp
      8947      9665: contig of 919 bp in length
      9666      9865: gap of 100 bp
      9867      10883: contig of 918 bp in length
      10884      10983: gap of 100 bp
      10984      11909: contig of 926 bp in length
      11910      12909: gap of 100 bp
      12910      12928: contig of 919 bp in length
      12929      13028: gap of 100 bp
      13029      13930: contig of 902 bp in length
      13931      14030: gap of 100 bp
      14031      14924: contig of 894 bp in length
      14925      15024: gap of 100 bp
      15025      15928: contig of 904 bp in length
      15929      16028: gap of 100 bp
      16029      16893: contig of 865 bp in length
      16894      16993: gap of 100 bp
      16994      17861: contig of 868 bp in length
      17862      17961: gap of 100 bp
      17962      18856: contig of 895 bp in length
      18857      18956: gap of 100 bp
      18957      19822: contig of 866 bp in length
      19823      19922: gap of 100 bp
      19923      20840: contig of 918 bp in length
      20841      20940: gap of 100 bp
      20941      21855: contig of 915 bp in length
      21856      21955: gap of 100 bp
      21956      22811: contig of 856 bp in length
      22812      22911: gap of 100 bp
      22912      23825: contig of 914 bp in length
      23826      23925: gap of 100 bp
      23926      24837: contig of 912 bp in length
      24838      24937: gap of 100 bp
      24938      25848: contig of 911 bp in length
      25849      25948: gap of 100 bp
      25949      26846: contig of 898 bp in length
      26847      26946: gap of 100 bp
      26947      27845: contig of 898 bp in length
      27845      27944: gap of 100 bp
      27945      28802: contig of 858 bp in length
      28803      28902: gap of 100 bp
      28903      29817: contig of 915 bp in length
      29817      29917: gap of 100 bp

```

```
* 29918 30805: contig of 888 bp in length
* 30905: gap of 100 bp
* 31906 31808: contig of 903 bp in length
* 31909 31808: gap of 100 bp
* 32905 32805: contig of 897 bp in length
* 32906 32805: gap of 100 bp
* 33767 33666: contig of 861 bp in length
* 33867 33762: contig of 896 bp in length
* 34763 34662: gap of 100 bp
* 35763 35662: contig of 900 bp in length
* 35863 35770: contig of 908 bp in length
* 36771 36870: gap of 100 bp
* 37832: contig of 962 bp in length
* 37933 37832: gap of 100 bp
* 38915 38814: contig of 882 bp in length
* 38915 37883: contig of 869 bp in length
* 39784 39883: gap of 100 bp
* 39884 40823: contig of 940 bp in length
* 40924 40823: gap of 100 bp
* 41824 41823: contig of 900 bp in length
* 41924 42808: contig of 885 bp in length
* 42909 42908: gap of 100 bp
* 43823 43823: contig of 914 bp in length
* 43923 43923: gap of 100 bp
* 44803 44803: contig of 887 bp in length
* 44910 44910: gap of 100 bp
* 45779 45779: contig of 870 bp in length
* 45880 45772: contig of 893 bp in length
* 46773 46872: gap of 100 bp
* 47759 47758: contig of 886 bp in length
* 47859 47858: gap of 100 bp
* 48779 48778: contig of 920 bp in length
* 48879 48779: gap of 100 bp
* 49741 49840: gap of 100 bp
* 49841 50694: contig of 854 bp in length
* 50695 50794: gap of 100 bp
* 50795 51699: contig of 905 bp in length
* 51700 51799: gap of 100 bp
* 51800 52685: contig of 886 bp in length
* 52686 52685: gap of 100 bp
* 52786 53703: contig of 920 bp in length
* 53706 53803: gap of 100 bp
* 53806 54718: contig of 913 bp in length
* 54719 54818: gap of 100 bp
* 54819 55719: contig of 901 bp in length
* 55720 55819: gap of 100 bp
* 55820 56733: contig of 914 bp in length
* 56734 56833: gap of 100 bp
* 56834 57742: contig of 909 bp in length
* 57743 57842: gap of 100 bp
* 57843 58753: contig of 911 bp in length
* 58754 58853: gap of 100 bp
* 58854 59748: contig of 895 bp in length
* 59749 59848: gap of 100 bp
* 59849 60721: contig of 872 bp in length
* 60721 61697: contig of 877 bp in length
* 61698 61997: gap of 100 bp
* 61998 62699: contig of 902 bp in length
* 62700 62799: gap of 100 bp
* 62800 63701: contig of 902 bp in length
* 63702 63801: gap of 100 bp
* 63802 64667: contig of 866 bp in length
* 64668 64767: gap of 100 bp
* 64768 65673: contig of 906 bp in length
* 65674 65773: gap of 100 bp
* 65774 66698: contig of 925 bp in length
```

```
* 66599 66798: gap of 100 bp
* 66799 67694: contig of 896 bp in length
* 67695 67794: gap of 100 bp
* 67795 68697: contig of 903 bp in length
* 68698 68797: gap of 100 bp
* 68798 69697: contig of 900 bp in length
* 69698 69797: gap of 100 bp
* 69798 70737: contig of 940 bp in length
* 70738 70837: gap of 100 bp
* 70838 71729: contig of 892 bp in length
* 71730 71829: gap of 100 bp
* 71830 72713: contig of 884 bp in length

Query Match 100.0%; Score 22; DB 2; Length 86794;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCTCTCATTTGG 22
    |||||
Db 13148 GATGCTGAAGTCTCTCATTTGG 13169

RESULT 7
CNS01RHY 164550 bp DNA linear PRI 26-APR-2001
LOCUS Human chromosome 14 DNA sequence BAC R-21008 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL162633
VERSION AL162633.3 GI:11545121
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 164550)
AUTHORS Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissenbach, J.
Sequencing of the human chromosome 14
Unpublished
REFERENCE 2 (bases 1 to 164550)
Genoscope.
Direct Submission
Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Dec 4, 2000 this sequence version replaced gi:7799785.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-930011
Downstream BAC (overlapping the SP6 end) : C-2184N20 (AC-AL359234)
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 16.67x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 :
1 - 9 : 1
10 - 19 : 8
20 - 29 : 19
30 - 39 : 125
40 - 49 : 742
50 - 59 : 1376
60 - 69 : 1816
70 - 79 : 4694
```

80 - 89 : 21302
90 - 99 : 134467

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

source
Location/Qualifiers
1. .164550
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-21008"
/clone_lib="RPCI-11"
142293.142420
/note="matching EMBL:H53462"
RDB:RH53794
dbSTS:STS6714
Identified using the e-PCR software (G. Schuler)"

STS

ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 164550;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTGG 22
|||||
Db 112337 GATGCTGAAGTCGTCTCATTGG 112358

RESULT 8

CNS05STC7 175988 bp DNA linear PRI 28-APR-2001
LOCUS Human chromosome 14 DNA sequence BAC R-930011 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION AL355073
VERSION AL355073.5 GI:13897293
KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

1 (bases 1 to 175988)
Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissbach, J.

Sequencing of the human chromosome 14

Unpublished

2 (bases 1 to 175988)

Genoscope.

Direct Submission

Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

On Apr 30, 2001 this sequence version replaced gi:12697154.

----- Genom Center

Center: Genoscope / Centre National de Sequencage

Center code: GS

Web site: http://www.genoscope.cns.fr/

Contact: Seqref@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : C-2588E21

Downstream BAC (overlapping the SP6 end) : R-21008 (AC-AL162633)

----- Summary Statistics

Assembly program: Phrap; version 2.0

Quality coverage: 7.58x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases

0 :

1 - 9 :

10 - 19 :

20 - 29 : 2
30 - 39 : 73
40 - 49 : 1732
50 - 59 : 6318
60 - 69 : 6755
70 - 79 : 14805
80 - 89 : 45266
90 - 99 : 101037

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

Location/Qualifiers
1. .175988
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-930011"
72485.72612
/note="matching EMBL:H53462"
RDB:RH53794
dbSTS:STS6714
Identified using the e-PCR software (G. Schuler)"

STS

ORIGIN

Query Match 100.0%; Score 22; DB 9; Length 175988;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTGG 22
|||||

Db 102561 GATGCTGAAGTCGTCTCATTGG 102540

RESULT 9

AC010097

LOCUS

DEFINITION 198278 bp DNA linear HTG 07-JUL-2000
Homo sapiens chromosome 14 clone RP11-378116, WORKING DRAFT
SEQUENCE, 18 unordered pieces.

ACCESSION AC010097

VERSION AC010097.6 GI:8954335

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 198278)

Waterston, R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 198278)

Waterston, R.H.

Direct Submission

Submitted (11-SEP-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Jul 7, 2000 this sequence version replaced gi:8568179.

COMMENT

----- Genom Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----

Center project name: H_NH0378116

----- Summary Statistics -----

Sequencing vector: p13; 88%

Chemistry: Dye-terminator ET; 73% of reads

Chemistry: Dye-terminator Big Dye; 27% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 191558 bases at least Q40

Consensus quality: 193468 bases at least Q30

Consensus quality: 194761 bases at least Q20

Insert size: 195000; agarose-fp
 Insert size: 19578; sum-of-contigs
 Quality coverage: 5.76 in Q20 bases; agarose-fp
 Quality coverage: 5.18 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1211: contig of 1211 bp in length
 * 1212: gap of unknown length
 * 1312: contig of 1635 bp in length
 * 2947: gap of unknown length
 * 3046: contig of 3190 bp in length
 * 3047: gap of unknown length
 * 6237: gap of unknown length
 * 6337: contig of 3027 bp in length
 * 9363: gap of unknown length
 * 9464: contig of 4532 bp in length
 * 13995: gap of unknown length
 * 14095: contig of 3853 bp in length
 * 17948: gap of unknown length
 * 18049: contig of 5217 bp in length
 * 23265: gap of unknown length
 * 23366: contig of 7083 bp in length
 * 30449: gap of unknown length
 * 30548: contig of 10960 bp in length
 * 30549: gap of unknown length
 * 41509: contig of 9894 bp in length
 * 41609: gap of unknown length
 * 51503: contig of 9714 bp in length
 * 51603: gap of unknown length
 * 61316: contig of 11531 bp in length
 * 61416: gap of unknown length
 * 72947: contig of 14280 bp in length
 * 72948: gap of unknown length
 * 73048: contig of 16344 bp in length
 * 87328: gap of unknown length
 * 87428: contig of 17071 bp in length
 * 103772: gap of unknown length
 * 120942: contig of 17833 bp in length
 * 121043: gap of unknown length
 * 138876: contig of 23646 bp in length
 * 138976: gap of unknown length
 * 162622: contig of 35557 bp in length.
 * 162722: Location/Qualifiers

FEATURES

source
 1. .198278
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone="RP11-378I16"

ORIGIN

Query Match 100.0%; Score 22; DB 2; Length 198278;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTGG 22
 |||||

Db 106124 GATGCTGAAGTCGTCCTCATTGG 106145
 |||||

RESULT 10

AC137805
 LOCUS AC137805 233753 bp DNA linear HTG 03-DEC-2002
 DEFINITION Homo sapiens chromosome 16 clone RP11-863C20, WORKING DRAFT
 ACCESSION AC137805
 VERSION AC137805.1 GI:26006547

KEYWORDS

SOURCE
 ORGANISM

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 Homo sapiens (human)

Homo sapiens
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 233753)
 DOE Joint Genome Institute.
 Sequencing of Human Chromosome 16
 Unpublished

2 (bases 1 to 233753)
 DOE Joint Genome Institute.
 Direct Submission

Submitted (03-DEC-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

Project Information
 Center Project Name: 1625964
 Center clone name: RPCI-11_863C20

Summary Statistics
 Consensus quality: 233513 bases at least Q40
 Consensus quality: 233587 bases at least Q30
 Consensus quality: 233607 bases at least Q20
 Estimated insert size: 160000; agarose-fp estimation
 Estimated insert size: 233653; sum-of-contigs estimation
 Quality coverage: 18.97 in Q20 bases; agarose-fp estimation
 Quality coverage: 12.99 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 23031: contig of 23031 bp in length
 * 23032: gap of unknown length
 * 23132: contig of 210622 bp in length.

Location/Qualifiers
 1. .233753
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-863C20"
 /clone_lib="RPCI human BAC library 11"

Query Match 100.0%; Score 22; DB 2; Length 233753;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTGG 22
 |||||

Db 131242 GATGCTGAAGTCGTCCTCATTGG 131263
 |||||

RESULT 11
 AK025208/c
 LOCUS AK025208 1890 bp mRNA linear PRI 13-SEP-2003
 DEFINITION Homo sapiens cDNA: FLJ21555 fis, clone COL06351.
 ACCESSION AK025208
 VERSION AK025208.1 GI:10437673
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
JOURNAL NEDO human cDNA sequencing project
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 1890)
 Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: fcdna@ms.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
FEATURES Location/Qualifiers
 source 1..1890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="COL06351"
 /tissue_type="colon"
 /clone_lib="COL"
 /note="Cloning vector pME18SFL3"
 238..633
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15084.1"
 /db_xref="GI:10437674"
 /translation="MSCHFKGLAPLPHVVTQGNCRFISCLGLTLPFFASSPFEVKVP VMSHRNIFQLKMSFTKKIKSGWSTLSIFLVRNLLIIGAAGSCMLQLSTFPLFG SAFSKGYSPPCRVRLDFLPLETVAQKN"
 CDS
 1..1890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="COL06351"
 /tissue_type="colon"
 /clone_lib="COL"
 /note="Cloning vector pME18SFL3"
 238..633
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15084.1"
 /db_xref="GI:10437674"
 /translation="MSCHFKGLAPLPHVVTQGNCRFISCLGLTLPFFASSPFEVKVP VMSHRNIFQLKMSFTKKIKSGWSTLSIFLVRNLLIIGAAGSCMLQLSTFPLFG SAFSKGYSPPCRVRLDFLPLETVAQKN"
 ORIGIN
 Query Match 92.7%; Score 20.4; DB 9; Length 1890;
 Best Local Similarity 95.5%; Pred. No. 18;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 GATGCTGAAGTCGCTCTCATTTG 22
 |||||
 1176 GACGCTGAAGTCGCTCTCATTTG 1355
 RESULT 12
 HANS07474 539 bp DNA linear PLN 19-JUN-2003
 LOCUS Heterobasidion insulare partial mnp1a gene for putative
 DEFINITION Mn-dependent peroxidase, exons 1-4, isolate B1279.
 ACCESSION AJ507474
 VERSION AJ507474.1 GI:23304315
 KEYWORDS Mn-dependent peroxidase; mnp1a gene.
 SOURCE Heterobasidion insulare
 ORGANISM Heterobasidion insulare
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Schizophyllaceae; Heterobasidion.
 REFERENCE 1
 AUTHORS Maijala,P.M., Harrington,T.C. and Raudaskoski,M.
 TITLE A peroxidase gene family and gene trees in Heterobasidion and related genera
 JOURNAL Mycologia 95 (2), 209-221 (2003)
 REFERENCE 2 (bases 1 to 539)
 AUTHORS Maijala,P.M.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2002) Maijala P.M., Applied Chemistry and Microbiology, University of Helsinki, P.O. Box 56, FIN-00014, FINLAND
FEATURES Location/Qualifiers
 source 1..539
 /organism="Heterobasidion insulare"
 /mol_type="genomic DNA"
 /isolate="B1279"
 /specific_host="Pinus coraiensis"
 /db_xref="taxon:135442"
 /country="China"
 1..539
 /gene="mnp1a"
 join(41..131,192..313,365..440,490..>539)
 /gene="mnp1a"
 /EC_number="1.11.1.13"
 /function="lignin degradation"
 /codon_start=1
 /product="putative Mn-dependent peroxidase"
 /protein_id="CAD47808.1"
 /db_xref="GI:23304316"
 /db_xref="GOA:Q8U1U2"
 /translation="GGADGSIIVPSDIETNFHANNGIDEIVVEOKPFIARHNITPGDF IQPAGAIGVSNCPGAPQLDFLGRPAPVAPADLTVPPEPDSVDSILARFNDTGFNA EVALLASHITIAADKVDVTIPGT"
 <1..131
 /gene="mnp1a"
 /number=1
 132..191
 /gene="mnp1a"
 /number=1
 192..313
 /gene="mnp1a"
 /number=2
 314..364
 /gene="mnp1a"
 /number=2
 365..440
 /gene="mnp1a"
 /number=3
 441..489
 /gene="mnp1a"
 /number=3
 490..>539
 /gene="mnp1a"
 /number=4
 ORIGIN
 Query Match 80.9%; Score 17.8; DB 8; Length 539;
 Best Local Similarity 90.5%; Pred. No. 4.1e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 1 GATGCTGAAGTCGCTCTCATTTG 21
 |||||
 412 GCTGCTGAAGTCGCTCGCATTTG 432
 RESULT 13
 HANS07473 549 bp DNA linear PLN 19-JUN-2003
 LOCUS Heterobasidion araucariae partial mnp1a gene for putative
 DEFINITION Mn-dependent peroxidase, exons 1-4, isolate B1080.
 ACCESSION AJ507473
 VERSION AJ507473.1 GI:23304313
 KEYWORDS Mn-dependent peroxidase; mnp1a gene.
 SOURCE Heterobasidion araucariae
 ORGANISM Heterobasidion araucariae
 Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Schizophyllaceae; Heterobasidion.
 REFERENCE 1
 AUTHORS Maijala,P.M., Harrington,T.C. and Raudaskoski,M.
 TITLE A peroxidase gene family and gene trees in Heterobasidion and related genera
 JOURNAL Mycologia 95 (2), 209-221 (2003)
 REFERENCE 2 (bases 1 to 549)
 AUTHORS Maijala,P.M.

Search completed: July 20, 2005, 18:01:22
Job time : 858.261 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 11:52:56 ; Search time 201.348 Seconds
(without alignments)
646.813 Million cell updates/sec

Title: US-10-041-030-7

Perfect score: 22

Sequence: 1 gatcgtgaagtcgtctcattgg 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	6	ABQ78313 Probe N63
2	22	100.0	60	6	ABN40168 Human spl
3	22	100.0	259	8	ABZ18999 Group IIX
4	22	100.0	554	6	ABT10524 Human pfe
5	22	100.0	864	4	AAH34231 Human/col
6	22	100.0	3614	4	AAI04355 Human rep
7	22	100.0	3615	4	AAI04354 Human rep
8	22	100.0	3752	4	AAI04353 Human rep
9	22	100.0	5579	8	ACC42349 Human MAP
10	17.2	78.2	2182	6	ABA95158 S. castel
11	17.2	78.2	2415	8	ACA45178 Prokaryot
12	17.2	78.2	3634	8	ABX77573 Prokaryot
13	17.2	78.2	4070	6	ABQ92052 Human pol
14	17.2	78.2	4130	2	AAV62749 Human sec
15	17.2	78.2	4799	12	ADG64976 Novel hum
16	17.2	78.2	110000	2	AAAX9190_01
17	17	77.3	958	4	AAI71176 Corynebaci
18	17	77.3	1581	5	AAH66619 C glutam
19	17	77.3	349980	5	AAH68529 C glutam
20	16.8	76.4	1770	8	ACA34475 Prokaryot

ALIGNMENTS

RESULT 1

ABQ78313

ID ABQ78313 standard; DNA; 22 BP.

XX ABQ78313;

AC ABQ78313;

DT 05-NOV-2002 (first entry)

XX Probe N63226QF to detect Pellino 2 gene in cancer cells.

DE Pellino 2; cancer; Pellino 1; cancer treatment; epithelial cancer;

XX Gastrointestinal tract cancer; probe; as.

OS Homo sapiens.

XX WO200259611-A2.

XX 01-AUG-2002.

XX 28-DEC-2001; 2001WO-US051368.

XX 02-JAN-2001; 2001US-0259502P.

XX (TULA-) TULARIK INC.

XX Powers S, Mu D, Xiang P, Peng Y;

XX WPI; 2002-619185/66.

XX Detecting cancer cells in mammalian sample, useful for identifying

XX inhibitors for treating cancer e.g. epithelial cancer, comprises

XX detecting an overexpression of, or increase in copy number of genes

XX encoding, Pellino 1 and Pellino 2.

XX Example 2; Page 55; 69pp; English.

XX Probes ABQ78313-15 were used to detect human Pellino 2 gene in cancer

XX cells. The specification describes a method for detecting cancer cells in

XX biological sample from a mammal. The method comprises detecting an

XX overexpression of, or increase in copy number of genes encoding,

XX polypeptides Pellino 1 or Pellino 2. The method is useful in detecting

XX cancer or propensity to develop cancer, monitoring the efficacy of cancer

XX treatment, identifying inhibitors of Pellino 1 and 2, inhibiting the

Ada02813 Mouse Egr
Adb72551 Mouse Egr
Adc85293 Mouse Egr
Adm74408 Murine ca
Continuation (16 o
Abac36278 Probe #14
Ad103150 DNA encod
Abac36278 Probe #14
Abac36278 Probe #48
Ad53383 Plant DNA
Abq70068 Listeria
Aah51151 Human ME1
Aas62696 cDNA sequ
Ad48878 Bacterial
Acf71856 Photorhab
Aca30358 Prokaryot
Adg63289 Novel hum
Abbs4378 cDNA enco
Ada53331 Human cod
Adr06611 Full leng
Abss67815 Human rec
Abz11699 Human pol
Adf82184 Leukaemia

CC expression and/or activity of Pellino 1 and 2 in cancer cells, and
 CC treating cancer or inhibiting proliferation of cancer. The cancer can be
 CC epithelial cancer, such as lung, colon, ovarian, breast, prostate,
 CC kidney, stomach, bladder, or any cancer of the gastrointestinal tract
 XX
 SQ Sequence 22 BP; 4 A; 4 C; 7 G; 7 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.39; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

OY 1 GATGCTGAAGTCGTCCTCATTGG 22
 |||||
 DB 1 GATGCTGAAGTCGTCCTCATTGG 22

RESULT 2
 ABN40168/c
 ID ABN40168 standard; DNA; 60 BP.
 XX
 AC ABN40168;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:12916.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 KW splice variant; transcriptome; oligonucleotide library; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-1B001903.
 XX
 PR 28-JUL-2000; 2000US-0221607P.
 PR 02-MAY-2001; 2001US-0287724P.
 XX
 PA (COMP-) COMPUGEN INC.
 XX
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 XX
 DR WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a
 PT genome, useful for detecting tissue-, pathology-, and developmental-
 PT specific genes.
 XX
 PS Example 1; SEQ ID NO 12916; 47pp; English.
 XX
 CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-
 CC)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises several
 CC oligonucleotides, each capable of hybridizing selectively to a set of
 CC messenger RNAs transcribed from a given transcription unit of the genome,
 CC which encodes one or more messenger RNA splice variants. The
 CC oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a particular
 CC biological or pathological state, and so allowing the detection of tissue
 CC - and pathology-specific genes such as those genes only expressed in
 CC specific tissue under a specific pathological condition; to detect
 CC developmental specific genes; and to detect RNA transcripts and splice
 CC variants of a transcriptome of a patient suffering from a particular
 CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
 CC rats, humans and mice, which are used in the exemplification of the

CC present invention. N.B. The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 60 BP; 11 A; 15 C; 15 G; 19 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 60;
 Best Local Similarity 100.0%; Pred. No. 0.44; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

OY 1 GATGCTGAAGTCGTCCTCATTGG 22
 |||||
 DB 60 GATGCTGAAGTCGTCCTCATTGG 39

RESULT 3
 ABZ18999
 ID ABZ18999 standard; cDNA; 259 BP.
 XX
 AC ABZ18999;
 XX
 DT 23-JAN-2003 (first entry)
 XX
 DE Group III cDNA cancer related clone SEQ ID NO:1425.
 XX
 KW Human; cancer; tumour; therapy; diagnosis; CT antigen; CP antigen;
 KW immune response; virology; immunology; microbiology; molecular biology;
 KW recombinant DNA technology; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200278516-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US010421.
 XX
 PR 30-MAR-2001; 2001US-0280255P.
 PR 28-AUG-2001; 2001US-0315563P.
 PR 09-JAN-2002; 2002US-0347313P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Wang T, Wang S, Bangur CS, Gaiger A;
 XX
 DR WPI; 2003-058387/05.

XX New immunogenic polynucleotides or polypeptides useful for diagnosing,
 PT preventing and treating cancer expressing CT or CP mRNA antigens, and in
 PT virology, immunology, microbiology, molecular biology and recombinant DNA
 PT techniques.
 XX
 PS Claim 1; SEQ ID NO 1425; 207pp; English.
 XX
 CC ABQ17575 to ABQ20506 represent isolated polynucleotide (I) sequences, and
 CC ABP54446 to ABP54472 represent protein (II) sequences, from the present
 CC invention. (I) and (II) have cytostatic activity and can be used in gene
 CC therapy and vaccines. (I), (II), antibodies and compositions from the
 CC present invention are useful for diagnosing, preventing and treating
 CC cancer, which expresses CT or CP mRNA antigens. They are useful for
 CC stimulating immune response. They can also be useful in virology,
 CC immunology, microbiology, molecular biology and recombinant DNA
 CC techniques. N.B. The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 259 BP; 89 A; 50 C; 59 G; 61 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 8; Length 259;
 Best Local Similarity 100.0%; Pred. No. 0.54; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0;

OY 1 GATGCTGAAGTCGTCCTCATTGG 22

Db 6 GATGCTGAAGTCGCTCTCATTTGG 27

|||||

RESULT 4

ABT10524

ID ABT10524 standard; cDNA; 554 BP.

XX AC ABT10524;

XX DT 04-DEC-2002 (first entry)

XX DE Human breast cancer associated coding sequence SEQ ID NO: 658.

XX DE Human; breast specific gene; breast cancer; differential expression;

XX KW cytotatic; gene therapy; gene; ss.

XX OS Homo sapiens.

XX PN WO200259271-A2.

XX PD 01-AUG-2002.

XX PF 25-JAN-2002; 2002WO-US002176.

XX PR 25-JAN-2001; 2001US-0263757P.

XX PR 25-APR-2001; 2001US-0286090P.

XX PR 23-MAY-2001; 2001US-0292517P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Orr MS, Nation M, Diggins JC, Zeng W;

XX WPI; 2002-674803/72.

XX Diagnosing breast cancer in a patient comprises detecting the level of

PT gene expression in cell or tissue samples, where a differential gene

PT expression is indicative of breast cancer.

XX Claim 1; SEQ ID NO 658; 260pp + Sequence Listing; English.

XX The present invention relates to methods of diagnosing breast cancer in a

CC patient, which comprise detecting the level of expression in a tissue

CC sample of two or more genes selected from those shown in ABT09867-

CC ABT1112, where a differential expression of the genes indicates breast

CC cancer. The methods are useful in diagnosing, treating, detecting the

CC progression, and in monitoring treatment of breast cancer in patients.

CC The methods are also useful as a screening tool for agents that modulate

CC the onset or progression of breast cancer. The breast cancer genes may be

CC used as diagnostic markers for the prediction or identification of the

CC malignant state of breast tissue, for confirming the type and progression

CC of cancer, and for drug screening and assays. The present sequence is a

CC coding sequence of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub.published_pct_sequences

XX SQ Sequence 554 BP; 196 A; 88 C; 92 G; 178 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 554;

Best Local Similarity 100.0%; Pred. No. 0.59;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCTCATTTGG 22

Db 446 GATGCTGAAGTCGCTCTCATTTGG 467

|||||

RESULT 5

AAH34231/C

ID AAH34231 standard; cDNA; 864 BP.

XX AC AAH34231;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen encoding cDNA SEQ ID NO:1313.

XX DE Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX KW colorectal carcinoma; ss.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US026524.

XX PR 29-SEP-1999; 99US-0157137P.

XX PR 03-NOV-1999; 99US-0163280P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX P-PSDB; AAG74826.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 1; Page 3049; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where the

CC proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene therapy

CC and vaccine production. N and P may be used in the prevention, diagnosis

CC and treatment of diseases associated with inappropriate P expression. For

CC example, N and P may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of P by expressing inactive proteins or to

CC supplement the patients own production of P. Additionally, N may be used

CC to produce the colon cancer-associated P, by inserting the nucleic acids

CC into a host cell and culturing the cell to express the proteins. N and P

CC can be used in the prevention, diagnosis and treatment of colorectal

CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent

CC sequences used in the exemplification of the present invention. N.B.

CC Pages 666 to 682 and page 7053 of the sequence listing were missing at

CC time of publication, meaning no sequences are present for SEQ ID NO:1027

CC to 1052, 7921 and 7922

XX SQ Sequence 864 BP; 247 A; 143 C; 157 G; 314 T; 0 U; 3 Other;

Query Match 100.0%; Score 22; DB 4; Length 864;

Best Local Similarity 100.0%; Pred. No. 0.63;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCTCATTTGG 22

Db 333 GATGCTGAAGTCGCTCTCATTTGG 312

|||||

RESULT 6

AAH04355/C

ID AAH04355 standard; DNA; 3614 BP.

XX AC AAH04355;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 7043.

XX KW Human; reproductive system related antigen; reproductive system disorder;

XX KW cancer; gene therapy; ds.

XX OS Homo sapiens.
XX PN WO200155320-A2.
XX XX
PD PF 02-AUG-2001.
XX XX
PF PF 17-JAN-2001; 2001WO-US001339.
XX XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
XX XX
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-465570/50.
 XX Isolated nucleic acid molecule encoding a reproductive system antigen is
 PT used in preventing, treating or ameliorating a medical condition.
 PT Disclosure; SEQ ID NO 7043; 1297pp + Sequence Listing; English.
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention
 XX
 SQ Sequence 3614 BP; 1063 A; 621 C; 657 G; 1273 T; 0 U; 0 Other;
 Query Match 100.0%; Score 22; DB 4; Length 3614;
 Best Local Similarity 100.0%; Pred. No. 0.76;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATGCTGAGTCTCTCATTTGG 22
 DB 3085 GATGCTGAGTCTCTCATTTGG 3064
 RESULT 7
 AAL04354/c
 ID AAL04354 standard; DNA; 3615 BP.
 XX AAL04354;
 AC
 DT 21-NOV-2001 (first entry)
 XX
 DE Human reproductive system related antigen DNA SEQ ID NO: 7042.
 XX
 KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ds.
 XX Homo sapiens.
 XX
 PN WO200155320-A2.
 PD
 XX 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001339.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 13-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 08-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.

PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234224P.
 PR 25-SEP-2000; 2000US-0234977P.
 PR 25-SEP-2000; 2000US-0234978P.
 PR 26-SEP-2000; 2000US-0234988P.
 PR 26-SEP-2000; 2000US-0234989P.
 PR 27-SEP-2000; 2000US-0234984P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 20-OCT-2000; 2000US-0241827P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-465570/50.
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen is
 PT used in preventing, treating or ameliorating a medical condition.
 XX
 XX Disclosure; SEQ ID NO 7041; 1297pp + Sequence Listing; English.
 XX
 XX The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders.
 CC including cancer. The present sequence is a genomic sequence encoding a
 CC protein of the invention
 XX
 XX Sequence 3752 BP; 1101 A; 648 C; 671 G; 1332 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 22; DB 4; Length 3752;
 Best Local Similarity 100.0%; Pred. NO. 0.76;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATGCTGAAGTGTCTTCATTGG 22
 |||||
 DB 3086 GATGCTGAAGTGTCTTCATTGG 3065
 RESULT 9
 ACC42349/C
 ID ACC42349 standard; cDNA; 5579 BP.
 XX
 AC ACC42349;
 XX
 DT 22-MAY-2003 (first entry)
 XX
 XX Human MAP kinase cascade activator #59 cDNA.
 XX
 KW Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
 KW antiinflammatory; immunomodulator; cytostatic; antiallergic; anti-HIV;
 KW antirheumatic; antiarthritic; antidiabetic; antidiabetic; gene therapy;
 KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
 KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 KW Iga nephritis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO2003008589-A1.
 XX
 XX 30-JAN-2003.
 XX
 PF 15-JUL-2002; 2002WO-JP007174.
 XX
 XX 18-JUL-2001; 2001JP-00218204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JAN-2002; 2002JP-00012176.
 XX
 XX (ASAH) ASAHI KASEI KOGYO KK.
 PA
 XX Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 PI
 XX WPI; 2003-229582/22.
 DR P-PSDB; ABR41083.
 XX
 XX Elki phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs

PT for e.g. inflammations, autoimmune diseases, viral diseases and cancer.
XX
PS Claim 4; Page 631-640; 762pp; Japanese.
XX
CC The invention relates to a novel purified protein having Elkl
CC phosphorylation activity and/or an activity of activating Elkl
CC phosphorylation kinase. A protein of the invention has antiinflammatory,
CC immunomodulator, virucide, cytostatic, antiallergic, antirheumatic,
CC antiarthritic, antidiabetic, antiasthmatic, and anti-HIV activity. The
CC polynucleotides may have a use in gene therapy. The gene and its encoded
CC protein are applicable in diagnosis of and developing drugs for e.g.
CC inflammations, autoimmune diseases, viral diseases and cancer such as
CC rheumatoid arthritis, diabetes; asthma, allergic rhinitis, AIDS, viral
CC hepatitis and IGA nephritis. The present sequence is used in the
CC exemplification of the invention
XX
SQ Sequence 5579 BP; 1535 A; 1111 C; 1214 G; 1719 T; 0 U; 0 Other;
Query Match 100.0%; Score 22; DB 8; Length 5579;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATGCTGAAGTCGTCCTCATGG 22
Db 5135 GATGCTGAAGTCGTCCTCATGG 5114
RESULT 10
ABA95158/c
ID ABA95158 standard; DNA; 2182 BP.
XX
AC ABA95158;
XX
DT 07-AUG-2003 (revised)
DT 20-MAY-2002 (first entry)
XX
DE S. castellii glucoamylase (GAM) promoter sequence.
XX
KW Yeast; glucoamylase; GAM gene; gene expression; transcription; promoter;
KW ds.
XX
OS Debaryomyces sp.
XX
FH Key Location/Qualifiers
FT TATA_signal 1561..1565
FT CAAT_signal 1571..1574
FT TATA_signal 1626..1643
FT CAAT_signal 1709..1712
FT TATA_signal 1730..1734
FT CAAT_signal 1776..1779
FT CAAT_signal 1805..1808
FT CAAT_signal 1816..1819
FT TATA_signal 1864..1868
FT TATA_signal 1884..1887
FT TATA_signal 1937..1943
FT CAAT_signal 1963..1966
FT CAAT_signal 2015..2018
FT TATA_signal 2034..2037
FT TATA_signal 2039..2043

FT TATA_signal 2081..2090
FT TATA_signal 2135..2139
FT misc_feature 2148..2150
FT /note= "start codon"
PN WO200212516-A2.
PD 14-FEB-2002.
XX
XX 03-AUG-2001; 2001WO-US024476.
XX 04-AUG-2000; 2000US-00632314.
XX 02-AUG-2001; 2001US-00921944.
XX (BATT) BATTELLE MEMORIAL INST.
XX Gao J, Skeen RS, Hooker BS, Anderson DB;
XX WPI; 2002-227159/28.
XX Novel isolated yeast promoter native to Schwannomyces castellii and
XX located upstream of and in control of glucoamylase gene, useful for
XX expressing gene of interest in bacterial, yeast, mold, plant/plant cell
XX species.
XX Example 2; Fig 4; 28pp; English.
XX
XX The invention provides an isolated yeast promoter, which is native to
XX Schwannomyces castellii (ATCC 26077) and located upstream of and in
XX control of a glucoamylase (GAM) gene. The GAM promoter or a vector
XX comprising the promoter is useful for expressing a gene of interest in
XX bacterial, yeast, mold and plant/plant cell species. The GAM promoter is
XX useful for regulating strong gene expression in starch culture medium,
XX for directing transcription or expression of a gene of interest, for
XX regulating native or foreign gene expression in native or heterologous
XX host strains with starch, and for certain types of metabolic pathway
XX controlling and foreign pathway accumulation. The present sequence
XX represents the S. castellii GAM promoter sequence. (Updated on 07-AUG-
XX 2003 to correct OS field.)
XX
SQ Sequence 2182 BP; 713 A; 355 C; 343 G; 771 T; 0 U; 0 Other;
Query Match 78.2%; Score 17.2; DB 6; Length 2182;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GATGCTGAAGTCGTCCTCATGG 22
Db 103 GAGGCTGAAGTCGTCCTCATGG 82
RESULT 11
ACA45178/c
ID ACA45178 standard; DNA; 2415 BP.
XX
XX ACA45178;
XX
XX 19-JUN-2003 (first entry)
XX
XX Prokaryotic essential gene #26835.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Pseudomonas syringae.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US0009107.
XX
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haeelbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR P-PSDB; ABU41308.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 33048; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2415 BP; 605 A; 699 C; 658 G; 453 T; 0 U; 0 Other;
SQ
Query Match 78.2%; Score 17.2; DB 8; Length 2415;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GATGCTGAAGTCGTCCTCATGG 22
|||||
Db 189 GATGCTGAAGTCGTCGCAATGG 168
RESULT 12
ABX77573
ID ABX77573 standard; cDNA; 3634 BP.
XX
XX AC ABX77573;
XX
XX 09-APR-2003 (first entry)

XX Differentially expressed breast cancer associated cDNA #68.
DE
XX
XX Breast cancer; differential gene expression; BC-cDNA;
KW breast cancer diagnosis; breast cancer monitoring;
KW breast cancer treatment; breast cancer staging; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX US2002156263-A1.
PN
XX 24-OCT-2002.
PD
XX 04-OCT-2001; 2001US-00974298.
PF
XX 05-OCT-2000; 2000US-0238331P.
PR
XX (CHEN/) CHEN H.
PA
XX Chen H;
PI
XX
XX WPI; 2003-182653/18.
DR
XX
XX New cDNAs, which are differentially expressed in (metastatic) breast
PT cancer useful for diagnosing or staging, breast cancer, or for monitoring
PT the treatment of breast cancer in an individual.
XX
XX Claim 1; SEQ ID NO 83; 30pp; English.
PS
XX The invention describes a combination of cDNAs (designated BC-cDNAs),
CC which are differentially expressed in breast cancer. The combination
CC includes 152 cDNA sequences, or their complements. The protein encoded by
CC any of these BC-cDNAs is useful for screening several molecules or
CC compounds to identify at least one ligand that specifically binds the
CC protein, producing or preparing polyclonal or monoclonal antibodies, or
CC purifying antibodies from a sample. The antibodies, which specifically
CC bind the protein differentially expressed in breast cancer is useful for
CC detecting the expression of a protein in a sample. The BC-cDNAs are also
CC useful for diagnosing, monitoring the treatment of, or staging, breast
CC cancer. This sequence represents a differentially expressed breast cancer
CC associated cDNA. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=20020156263
XX
XX Sequence 3634 BP; 992 A; 928 G; 932 C; 781 T; 0 U; 1 Other;
SQ
Query Match 78.2%; Score 17.2; DB 8; Length 3634;
Best Local Similarity 86.4%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GATGCTGAAGTCGTCCTCATGG 22
|||||
Db 805 GATGCTGAAGTCCTCACATGG 826
RESULT 13
ABQ92052
ID ABQ92052 standard; cDNA; 4070 BP.
XX
XX AC ABQ92052;
XX
XX 04-OCT-2002 (first entry)
DT
XX Human polynucleotide SEQ ID NO 49.
DE
XX Human; cytostatic; antirheumatic; antiarthritic; vulnery; analgesic;
KW antinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
KW antiulcer; fungicide; antidiabetic; antisthmatic; antiallergic;
KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
KW stem cell; growth factor; nervous system disease; neuropathy;

KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX Homo sapiens.
 OS US2002065394-A1.
 PN 30-MAY-2002.
 PD 22-DEC-2000; 2000US-00745763.
 XX 18-MAR-1998; 98US-00040963.
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LAVA/) LAVALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREAC/) TREACY M.
 PA (SPAUL/) SPAULDING V.
 XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 XX WPI; 2002-582343/62.
 DR P-PSDB; ABP61838.
 XX Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.
 XX Claim 174; Page 182-183; 284pp; English.
 PS The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumors. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention
 XX Sequence 4070 BP; 1098 A; 1034 C; 1012 G; 921 T; 0 U; 5 Other;

Query Match 78.2%; Score 17.2; DB 6; Length 4070;
 Best Local Similarity 86.4%; Pred. No. 2.1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTCG 22
 DB 1357 GATGCTGAAGTCCTCACAGTCG 1378
 RESULT 14
 AAV62749
 ID AAV62749 standard; cDNA; 4130 BP.
 XX AAV62749;
 XX 15-FEB-1999 (first entry)
 XX Human secreted protein clone er418_5 cDNA.
 DE Secreted protein; human; er418_5; ds.
 KW Homo sapiens.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 571..3306
 FT /*tag= a
 XX WO9846757-A2.
 PN 22-OCT-1998.
 PD 14-APR-1998; 98WO-US007999.
 XX 15-APR-1997; 97US-00843374.
 PR 13-APR-1998; 98US-00059487.
 XX (GEWY) GENETICS INST INC.
 XX Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino MJ;
 XX WPI; 1998-569731/48.
 DR P-PSDB; AAW74723.
 XX New polynucleotide(s) encoding secreted human proteins - are derived
 PT from, e.g. human foetal brain or foetal kidney cDNA libraries,
 PT potentially useful as, e.g. vaccines or thrombolytic agents.
 XX Claim 19; Page 76-78; 120pp; English.
 PS Full-length cDNA clone er418_5 includes an open reading frame encoding a
 CC human secreted protein (see AAW74723). It was isolated from a human
 CC foetal brain cDNA library using methods which are selective for cDNAs
 CC encoding secreted proteins, or was identified as encoding a secreted or
 CC transmembrane protein on the basis of computer analysis of the amino acid
 CC sequence of the encoded protein. It shows some similarity to database
 CC sequences. The invention provides polynucleotides (see AAV62746-55) from
 CC human foetal brain, adult testis, adult brain, adult kidney and foetal
 CC kidney (all deposited as composite clone ATCC 98404), which encode human
 CC secreted proteins (see AAW74720-29). The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data are given. Suggested activities
 CC include nutritional activity, immune stimulating (e.g. as vaccines) or
 CC suppressing activity, haematopoiesis regulating activity, tissue growth
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
 CC haemostatic and thrombolytic activity, receptor/ligand activity,
 CC antiinflammatory activity, cadherin/tumour invasion suppressor activity,
 CC and tumour inhibition activity. The polynucleotides are also stated to be
 CC useful for gene therapy, and can be used in recombinant production of the
 CC polypeptides
 XX Sequence 4130 BP; 1113 A; 1051 C; 1024 G; 936 T; 0 U; 6 Other;

Query Match 78.2%; Score 17.2; DB 2; Length 4130;
 Best Local Similarity 86.4%; Pred. No. 2.1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCTCATTTG 22
 |||||
 Db 1417 GATGCTGAAGTCGCTCTCACAGTGG 1438

RESULT 15

ADQ64976
 ID ADQ64976 standard; cDNA; 4799 BP.
 XX
 AC ADQ64976;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Novel human cDNA sequence #2137.
 XX
 KW ss; gene; osteopathic; neuroprotective; neurotropic; antiparkinsonian;
 KW cytotatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
 KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
 KW cancer.
 XX
 OS Homo sapiens.
 XX
 PN EP1440981-A2.
 XX
 PD 28-JUL-2004.
 XX
 PF 21-JAN-2004; 2004EP-00001196.
 XX
 PR 21-JAN-2003; 2003JP-00102206.
 XX
 PR 09-MAY-2003; 2003JP-00131392.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Nagai K, Irie R;
 XX
 DR WPI; 2004-535376/52.
 DR P-PSDB; ADQ67164.
 XX
 PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 PS Claim 1; SEQ ID NO 2137; 2449pp; English.
 XX
 CC The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.
 XX
 SQ Sequence 4799 BP; 1251 A; 1234 C; 1233 G; 1081 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 12; Length 4799;
 Best Local Similarity 86.4%; Pred. No. 2.1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCTCATTTG 22
 |||||
 Db 2101 GATGCTGAAGTCGCTCTCACAGTGG 2122

Search completed: July 20, 2005, 17:31:34
 Job time : 208.348 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 16:53:52 ; Search time 69.3478 Seconds
(without alignments)
519.095 Million cell updates/sec

Title: US-10-041-030-7

Perfect score: 22

Sequence: 1 gatgctgaagtcgtctcattgg 22

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfileseq.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17.2	78.2	579	4	US-09-248-796A-12736
2	17.2	78.2	1230025	4	US-09-198-452A-1
3	17.2	78.2	1230230	4	US-09-438-185A-1
4	17	77.3	958	4	US-09-602-777A-77
5	16.8	76.4	1830121	4	US-09-557-884-1
6	16.8	76.4	1830121	4	US-09-643-990A-1
7	16.4	74.5	601	4	US-09-949-016-199193
8	16.4	74.5	98708	4	US-09-949-016-16392
9	16.2	73.6	204	4	US-09-540-236-836
10	16.2	73.6	318	4	US-09-248-796A-8067
11	16.2	73.6	601	4	US-09-949-016-187670
12	16.2	73.6	1001	4	US-09-671-317-42
13	16.2	73.6	6553	4	US-09-799-451-581
14	16.2	73.6	49617	4	US-09-596-002-28
15	16.2	73.6	66055	4	US-09-949-016-13292
16	16.2	73.6	96690	4	US-09-949-016-17103
17	16.2	73.6	194915	4	US-09-949-016-15584
18	16.2	73.6	390890	4	US-09-949-016-14720
19	15.8	71.8	601	4	US-09-949-016-91218
20	15.8	71.8	601	4	US-09-949-016-163448
21	15.8	71.8	3307	4	US-09-949-016-1973
22	15.8	71.8	3412	4	US-09-949-016-2202
23	15.8	71.8	3415	4	US-09-949-016-693
24	15.8	71.8	3418	3	US-09-193-562B-29
25	15.8	71.8	3418	4	US-10-055-412B-29
26	15.8	71.8	25111	4	US-09-949-016-12435
27	15.8	71.8	25111	4	US-09-949-016-13944

28	15.8	71.8	85869	4	US-09-949-016-12017	Sequence 12017, A
29	15.8	71.8	85878	4	US-09-949-016-16321	Sequence 16321, A
30	15.8	71.8	110266	4	US-09-949-016-14913	Sequence 14913, A
31	15.8	71.8	110266	4	US-09-949-016-14914	Sequence 14914, A
32	15.8	71.8	110266	4	US-09-949-016-14915	Sequence 14915, A
33	15.8	71.8	110266	4	US-09-949-016-14916	Sequence 14916, A
34	15.8	71.8	110266	4	US-09-949-016-14917	Sequence 14917, A
35	15.8	71.8	110266	4	US-09-949-016-14918	Sequence 14918, A
36	15.8	71.8	110266	4	US-09-949-016-14919	Sequence 14919, A
37	15.8	71.8	110266	4	US-09-949-016-14920	Sequence 14920, A
38	15.8	71.8	110266	4	US-09-949-016-14921	Sequence 14921, A
39	15.8	71.8	110266	4	US-09-949-016-14922	Sequence 14922, A
40	15.8	71.8	125902	4	US-09-949-016-13715	Sequence 13715, A
41	15.8	71.8	175236	4	US-09-949-016-14353	Sequence 14353, A
42	15.6	70.9	287	1	US-09-985-799-70	Sequence 70, Appl
43	15.6	70.9	287	1	US-09-977-371-70	Sequence 70, Appl
44	15.6	70.9	287	1	US-08-594-031-70	Sequence 70, Appl
45	15.6	70.9	413	4	US-09-733-685-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-248-796A-12736
; Sequence 12736, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 12736
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-12736

Query Match 78.2%; Score 17.2; DB 4; Length 579;
Best Local Similarity 86.4%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GATGCTGAAGTCGTCTCATTGG 22
DB 511 GTTGCTGAAGCCTTCTCATTGG 532

RESULT 2
US-09-198-452A-1
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc_feature

; PRIOR FILING DATE: 1999-09-03
 ; NUMBER OF SEQ ID NOS: 442
 ; SEQ ID NO 77
 ; LENGTH: 958
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(958)
 ; OTHER INFORMATION: RXA01151
 US-09-602-777A-77

Query Match 77.3%; Score 17; DB 4; Length 958;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTC 17
 |||||
 Db 241 GATGCTGAAGTCGTCTC 257

RESULT 5

US-09-557-884-1/c
 ; Sequence 1, Application US/09557884
 ; Patent No. 6506581
 ; GENERAL INFORMATION:
 APPLICANT: Fleischmann et al.
 TITLE OF INVENTION: The Nucleotide sequence of
 the Haemophilus influenzae Rd Genome, Fragments
 Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS v6.22
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/557,884
 FILING DATE: 25-Apr-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/476,102
 FILING DATE: JUN-5-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Michelle S. Marks
 REGISTRATION NUMBER: 41,971
 REFERENCE/DOCKET NUMBER: PB186P3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8504
 TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-557-884-1

Query Match 76.4%; Score 16.8; DB 4; Length 1830121;
 Best Local Similarity 90.0%; Pred. No. 2.6e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATGCTGAAGTCGTCTCATTTG 21
 |||||

Db 1553535 ATGCTGAAGTCGTCTCATCG 1553516

RESULT 6

US-09-643-990A-1/c
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289
 ; GENERAL INFORMATION:
 APPLICANT: Robert D. Fleischmann
 Mark D. Adams
 Owen White
 Hamilton O. Smith
 J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
 the Haemophilus influenzae Rd Genome, Fragments
 Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville,
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS v6.22
 SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/643,990A
 FILING DATE: 23-Aug-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,429
 FILING DATE: 1995-06-07
 APPLICATION NUMBER: 08/426,787
 FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:
 NAME: Kenley K. Hoover
 REGISTRATION NUMBER: 40,302
 REFERENCE/DOCKET NUMBER: PB186P1C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-610-5790
 TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-643-990A-1

Query Match 76.4%; Score 16.8; DB 4; Length 1830121;
 Best Local Similarity 90.0%; Pred. No. 2.6e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATGCTGAAGTCGTCTCATTTG 21
 |||||

Db 1553535 ATGCTGAAGTCGTCTCATCG 1553516

RESULT 7

US-09-949-016-199193/c
 ; Sequence 199193, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CLO01307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199193
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-199193

Query Match 74.5%; Score 16.4; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 97;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 TGCTGAAGTCGTCTCATGG 22
||:|||||
Db 303 TGSTGAGTCTTCTCATGG 284

RESULT 8

US-09-949-016-16392
; Sequence 16392, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16392
; LENGTH: 98708
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(98708)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16392

Query Match 74.5%; Score 16.4; DB 4; Length 98708;
Best Local Similarity 94.4%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ATGCTGAAGTCGTCTCAT 19
|||||
Db 5315 ATGCTGAAGTTGTCTAT 5332

RESULT 9

US-09-540-236-836
; Sequence 836, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 836
; LENGTH: 204
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-836

Query Match 73.6%; Score 16.2; DB 4; Length 204;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTTG 21
||:|||||
Db 172 GAAGCTGAAGTCGCCACATTTG 192

RESULT 10

US-09-248-796A-8067
; Sequence 8067, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 8067
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-8067

Query Match 73.6%; Score 16.2; DB 4; Length 318;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ATGCTGAAGTCGTCTCATTTG 22
|||||
Db 17 ACGCTGAAGTCGCTCATTTG 37

RESULT 11

US-09-949-016-187670
; Sequence 187670, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187670
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-187670

Query Match 73.6%; Score 16.2; DB 4; Length 601;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ATGCTGAAGTCGTCTCATTTG 22
 ||||| ||||| ||||| |||||
 Db 164 ATGCTGAAGTCGTCTCATTTG 184

RESULT 12

US-09-671-317-42
 ; Sequence 42, Application US/09671317
 ; Patent No. 6528260
 ; GENERAL INFORMATION:
 ; APPLICANT: Blumenfeld, Marta
 ; APPLICANT: Chumakov, Ilya
 ; APPLICANT: Bougueleret, Lydie
 ; APPLICANT: Cohen, Annick
 ; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
 ; FILE REFERENCE: 62 US3.CIP
 ; CURRENT APPLICATION NUMBER: US/09/671,317
 ; CURRENT FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US 09/536,178
 ; PRIOR FILING DATE: 2000-03-23
 ; PRIOR APPLICATION NUMBER: PCT/IB00/00403
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: US 60/126,269
 ; PRIOR FILING DATE: 1999-03-25
 ; PRIOR APPLICATION NUMBER: US 60/131,961
 ; PRIOR FILING DATE: 1999-04-30
 ; NUMBER OF SEQ ID NOS: 977
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 42
 ; LENGTH: 1001
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: allele
 ; LOCATION: 503
 ; OTHER INFORMATION: 12-723-293 : polymorphic base C or T
 ; NAME/KEY: misc_binding
 ; LOCATION: 484..502
 ; OTHER INFORMATION: 12-723-293.mis1
 ; NAME/KEY: misc_binding
 ; LOCATION: 504..523
 ; OTHER INFORMATION: 12-723-293.mis2, potential complement
 ; NAME/KEY: primer_bind
 ; LOCATION: 210..230
 ; OTHER INFORMATION: upstream amplification primer
 ; NAME/KEY: primer_bind
 ; LOCATION: 591..610
 ; OTHER INFORMATION: downstream amplification primer, complement
 ; NAME/KEY: misc_binding
 ; LOCATION: 491..515
 ; OTHER INFORMATION: 12-723-293 potential probe
 ; NAME/KEY: misc_feature
 ; LOCATION: 312..379
 ; OTHER INFORMATION: n=a, g, c or t
 ; US-09-671-317-42

Query Match 73.6%; Score 16.2; DB 4; Length 1001;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ATGCTGAAGTCGTCTCATTTG 22
 ||||| ||||| ||||| |||||
 Db 551 ATGTTGATGCTCTTCATTTG 571

RESULT 13

US-09-799-451-581/c
 ; Sequence 581, Application US/09799451
 ; Patent No. 6783969
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Ma, Yungqing
 ; APPLICANT: Yamazaki, Victoria
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Ghosh, Reena
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 803
 ; CURRENT APPLICATION NUMBER: US/09/799,451
 ; CURRENT FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 948
 ; SOFTWARE: pt_FL_genes Version 2.0
 ; SEQ ID NO 581
 ; LENGTH: 6553
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (272)..(5242)
 ; US-09-799-451-581

Query Match 73.6%; Score 16.2; DB 4; Length 6553;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ATGCTGAAGTCGTCTCATTTG 22
 ||||| ||||| ||||| |||||
 Db 2809 ATTATGAAGTCGTCTCATTTG 2789

RESULT 14

US-09-596-002-28
 ; Sequence 28, Application US/09596002
 ; Patent No. 6632836
 ; GENERAL INFORMATION:
 ; APPLICANT: Lagace, Robert, E.
 ; APPLICANT: Patterson, Chandra
 ; APPLICANT: Berg, Kim, L.
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
 ; FILE REFERENCE: PM-0008-4 US
 ; CURRENT APPLICATION NUMBER: US/09/596,002
 ; CURRENT FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: 60/140,121
 ; PRIOR FILING DATE: 1999-06-18
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 28
 ; LENGTH: 49617
 ; TYPE: DNA
 ; ORGANISM: M. catarrhalis
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: incyte template ID No. 6632636 28
 ; PUBLICATION INFORMATION:
 ; US-09-596-002-28

Query Match 73.6%; Score 16.2; DB 4; Length 49617;
 Best Local Similarity 85.7%; Pred. No. 3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCTCATTTG 21

Db 22308 GAAGCTGAAGTCGCCACATTG 22328
|| ||||| ||||| |||||

RESULT 15
US-09-949-016-13292
; Sequence 13292, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13292
; LENGTH: 66065
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(66065)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13292

Query Match 73.6%; Score 16.2; DB 4; Length 66065;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GATGCTGAAGTCGCTCTCATTG 21
||| ||||| ||||| ||||| |||||
Db 63263 GATGCTGATGCTCTCTATTG 63283

Search completed: July 20, 2005, 18:58:55
Job time : 79.3478 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 18:01:31 ; Search time 308.957 Seconds
(without alignments)
451.758 Million cell updates/sec

Title: US-10-041-030-7
Perfect score: 22
Sequence: 1 gatcgtgaagtcgtctcattgg 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7173243 seqs, 3172129809 residues
Total number of hits satisfying chosen parameters: 14346486

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	13	US-10-041-030-7
2	22	100.0	60	10	US-09-908-975-12916
3	22	100.0	864	15	US-10-106-698-1323
4	22	100.0	3614	10	US-09-764-891-7043
5	22	100.0	3615	10	US-09-764-891-7042
6	22	100.0	3752	10	US-09-764-891-7041
7	22	100.0	5579	14	US-10-197-666A-135

RESULT 1
US-10-041-030-7
; Sequence 7, Application US/10041030
; Publication No. US20020150934A1
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; TITLE OF INVENTION: Pellino Polypeptides and Polynucleotides
; FILE REFERENCE: 018781-006810US
; CURRENT APPLICATION NUMBER: US/10/041,030
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: oligonucleotide
; OTHER INFORMATION: probe N63226QF to pellino 2' untranslated region
US-10-041-030-7
Query Match 100.0%; Score 22; DB 13; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.67;

ALIGNMENTS

c	8	22	100.0	5597	17	US-10-085-117-287	Sequence 287, App
c	9	22	100.0	5921	13	US-10-041-030-3	Sequence 3, Appli
c	10	22	100.0	202802	17	US-10-085-117-286	Sequence 286, App
c	11	17.4	79.1	498	20	US-10-425-115-81792	Sequence 81792, A
c	12	17.4	79.1	593	13	US-10-027-632-46999	Sequence 46999, A
c	13	17.4	79.1	593	17	US-10-027-632-46999	Sequence 46999, A
c	14	17.2	78.2	1154	18	US-10-424-599-97451	Sequence 97451, A
c	15	17.2	78.2	2182	9	US-09-921-9448-8	Sequence 8, Appli
c	16	17.2	78.2	2415	17	US-10-282-122A-33048	Sequence 33048, A
c	17	17.2	78.2	3634	9	US-09-974-298-83	Sequence 93, Appli
c	18	17.2	78.2	4130	9	US-09-745-763-139	Sequence 139, App
c	19	17.2	78.2	1230025	17	US-10-289-762-1	Sequence 1, Appli
c	20	17	77.3	1581	9	US-09-738-626-1654	Sequence 1654, Ap
c	21	17	77.3	1980090	20	US-10-719-993-6815	Sequence 6815, Ap
c	22	17	77.3	1980090	21	US-10-741-600-17676	Sequence 17676, A
c	23	17	77.3	3309400	9	US-09-738-626-1	Sequence 1, Appli
c	24	16.8	76.4	1770	17	US-10-282-122A-22345	Sequence 22345, A
c	25	16.8	76.4	2065	18	US-10-425-114-8055	Sequence 8055, Ap
c	26	16.8	76.4	2097	18	US-10-424-599-22327	Sequence 22327, A
c	27	16.8	76.4	24080	11	US-09-997-722-79	Sequence 79, Appli
c	28	16.8	76.4	1830121	17	US-10-329-670-1	Sequence 1, Appli
c	29	16.8	76.4	1830121	20	US-10-158-865-1	Sequence 1, Appli
c	30	16.8	76.4	1830121	22	US-10-981-887-1	Sequence 1, Appli
c	31	16.4	74.5	639	13	US-10-027-632-5793	Sequence 5793, Ap
c	32	16.4	74.5	639	13	US-10-027-632-5794	Sequence 5794, Ap
c	33	16.4	74.5	639	17	US-10-027-632-5793	Sequence 5793, Ap
c	34	16.4	74.5	639	17	US-10-027-632-5794	Sequence 5794, Ap
c	35	16.4	74.5	648	13	US-10-027-632-5780	Sequence 5780, A
c	36	16.4	74.5	648	13	US-10-027-632-5781	Sequence 5781, A
c	37	16.4	74.5	648	13	US-10-027-632-67920	Sequence 67920, A
c	38	16.4	74.5	648	17	US-10-027-632-57780	Sequence 57780, A
c	39	16.4	74.5	648	17	US-10-027-632-57781	Sequence 57781, A
c	40	16.4	74.5	648	17	US-10-027-632-67920	Sequence 67920, A
c	41	16.4	74.5	1126	18	US-10-424-599-131822	Sequence 131822, A
c	42	16.4	74.5	24902	13	US-10-087-192-1045	Sequence 1045, Ap
c	43	16.2	73.6	160	9	US-09-864-761-21598	Sequence 21598, A
c	44	16.2	73.6	245	18	US-10-424-599-29593	Sequence 29593, A
c	45	16.2	73.6	454	9	US-09-783-590-4446	Sequence 4446, Ap

```
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCCTCATTGG 22
   |||||
Db 1 GATGCTGAAGTCGTCCTCATTGG 22

RESULT 2
US-09-908-975-12916/c
; Sequence 12916, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12916
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-12916

Query Match 100.0%; Score 22; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCCTCATTGG 22
   |||||
Db 60 GATGCTGAAGTCGTCCTCATTGG 39

RESULT 3
US-10-106-698-1323/c
; Sequence 1323, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1323
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (815)..(815)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (848)..(848)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
```

```
; LOCATION: (862)..(862)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1323

Query Match 100.0%; Score 22; DB 15; Length 864;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCCTCATTGG 22
   |||||
Db 333 GATGCTGAAGTCGTCCTCATTGG 312

RESULT 4
US-09-764-891-7043/c
; Sequence 7043, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7043
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7043

Query Match 100.0%; Score 22; DB 10; Length 3614;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCCTCATTGG 22
   |||||
Db 3085 GATGCTGAAGTCGTCCTCATTGG 3064

RESULT 5
US-09-764-891-7042/c
; Sequence 7042, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7042
; LENGTH: 3615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7042

Query Match 100.0%; Score 22; DB 10; Length 3615;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGAAGTCGTCCTCATTGG 22
   |||||
Db 3086 GATGCTGAAGTCGTCCTCATTGG 3065

RESULT 6
US-09-764-891-7041/c
; Sequence 7041, Application US/09764891
; Publication No. US20030077808A1
```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7041
; LENGTH: 3752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7041

Query Match 100.0%; Score 22; DB 10; Length 3752;
Best Local Similarity 100.0%; Pred. No. 0.97; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTTGG 22
|||||
Db 3086 GATGCTGAAGTCGTCCTCATTTGG 3065

RESULT 7

US-10-197-666A-135/c
; Sequence 135, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:

; APPLICANT: ASAH KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Elki phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 135
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (177)..(1436)
US-10-197-666A-135

Query Match 100.0%; Score 22; DB 14; Length 5579;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTTGG 22
|||||
Db 5135 GATGCTGAAGTCGTCCTCATTTGG 5114

RESULT 8

US-10-085-117-287/c
; Sequence 287, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:

; APPLICANT: Engelhard, David W.
; APPLICANT: Morris, Eric K.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287
; LENGTH: 5597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-287

Query Match 100.0%; Score 22; DB 17; Length 5597;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTTGG 22
|||||
Db 5135 GATGCTGAAGTCGTCCTCATTTGG 5114

RESULT 9

US-10-041-030-3/c
; Sequence 3, Application US/10041030
; Publication No. US20020150934A1
; GENERAL INFORMATION:

; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tuldrak Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; FILE REFERENCE: 018781-006810US
; CURRENT APPLICATION NUMBER: US/10/041,030
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5921
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1763)
; OTHER INFORMATION: human pellino 2
US-10-041-030-3

Query Match 100.0%; Score 22; DB 13; Length 5921;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATTTGG 22
|||||
Db 5459 GATGCTGAAGTCGTCCTCATTTGG 5438

RESULT 10

US-10-085-117-286/c
; Sequence 286, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586

```
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 202802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-286

Query Match      100.0%; Score 22; DB 17; Length 202802;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GATGCTGAAGTCGTCCTCATGG 22
Db      192359 GATGCTGAAGTCGTCCTCATGG 192338

RESULT 11
US-10-425-115-81792
; Sequence 81792, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81792
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174609C.1
US-10-425-115-81792

Query Match      79.1%; Score 17.4; DB 20; Length 498;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 ATGCTGAAGTCGTCCTCATTT 20
Db      8 ATGTTGAAGTCGTCCTCATTT 26

RESULT 12
US-10-027-632-46999/c
; Sequence 46999, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46999
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-46999

Query Match      79.1%; Score 17.4; DB 17; Length 593;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GATGCTGAAGTCGTCCTCAT 19
Db      230 GATGCTGAAGTTGTCTCAT 212

RESULT 13
US-10-027-632-46999/c
; Sequence 46999, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46999
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-46999

Query Match      79.1%; Score 17.4; DB 17; Length 593;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GATGCTGAAGTCGTCCTCAT 19
Db      230 GATGCTGAAGTTGTCTCAT 212

RESULT 14
US-10-424-599-97451/c
; Sequence 97451, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
```

0000000000

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 15:43:06 ; Search time 1573.48 Seconds
(without alignments)
532.206 Million cell updates/sec

Title: US-10-041-030-7

Perfect score: 22

Sequence: 1 gatgctgaagtcgtctcattgg 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_gsal:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	22	100.0	403	2	AW386118 RC3-PT000
C 2	22	100.0	451	1	AA905101 OK09206.8
C 3	22	100.0	493	2	AW173368 xj78f06.x
C 4	22	100.0	514	1	AI804454 tc71e02.x
C 5	22	100.0	531	1	AI082809 ox78a02.x
C 6	22	100.0	554	1	AI768516 wh22g11.x
C 7	22	100.0	555	2	BE466421 hz21e08.x
C 8	22	100.0	654	7	CN308828 170006001
C 9	22	100.0	687	5	BQ016066 UI-H-DT1-
C 10	22	100.0	782	5	BQ002692 UI-H-E11-
C 11	22	100.0	895	4	BG283778 602408077
C 12	22	100.0	901	2	BE786617 601478845
C 13	22	100.0	901	4	BG546951 602573943
C 14	22	100.0	943	4	BG388221 602413214
C 15	22	100.0	4308	3	HSMB01459 Homo sapi
C 16	20.4	92.7	171	6	CB117270 K-EST0162
C 17	18.8	85.5	645	5	BQ834924 Po ad 05B
C 18	18.6	84.5	1180	5	BX334080 BX334080
C 19	18.4	83.6	578	8	BH707520 BOHV803TF
C 20	17.8	80.9	324	7	CF933926 T-EST-B16
C 21	17.8	80.9	357	5	BQ488667 91-E8463-
C 22	17.8	80.9	378	8	B32701 HS-1015-B2-
C 23	17.8	80.9	438	6	BY688194 BY688194
C 24	17.8	80.9	440	9	TA168G07P T. brucei

C 25	17.8	80.9	553	9	TA286C03P
C 26	17.8	80.9	591	8	AO652813 T. brucei
C 27	17.8	80.9	600	8	AZ215044 Sheared D
C 28	17.8	80.9	649	2	AW318752 un05h01.y
C 29	17.8	80.9	649	5	BQ584470 E011861-0
C 30	17.8	80.9	724	7	CF866553 tric007xb
C 31	17.8	80.9	767	8	BH402703 AG-ND-119
C 32	17.8	80.9	773	7	CF681267 CCAIC82TR
C 33	17.8	80.9	803	6	CB896532 tric007xb
C 34	17.8	80.9	806	7	CF707757 CCAGY25TR
C 35	17.8	80.9	831	6	CB900843 tric024x1
C 36	17.8	80.9	831	7	CF870649 tric024x1
C 37	17.4	79.1	478	2	BE841995 MR4-ST009
C 38	17.4	79.1	490	2	AW581806 MR4-ST009
C 39	17.4	79.1	492	2	AW807650 MR4-ST009
C 40	17.4	79.1	636	8	AQ499367 HS_5199_B
C 41	17.4	79.1	673	5	BP152234 BP152234
C 42	17.4	79.1	973	4	BG709561 602673586
C 43	17.2	78.2	212	7	CK817477 hag8009xk
C 44	17.2	78.2	275	2	AW865771 QV3-SN002
C 45	17.2	78.2	322	1	AA319302 EST21987

ALIGNMENTS

RESULT 1
AW386118/c
LOCUS
RC3-PT0003-241199-011-d10 PT0003 Homo sapiens cDNA, mRNA linear EST 04-FEB-2000
ACCESSION
AW386118
VERSION
AW386118.1 GI:6890777
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 403)
HCGP <http://www.ludwig.org.br/ORESTES>.
AUTHORS
The FAPESP/LICR Human Cancer Genome Project
TITLE
Unpublished (1999)
JOURNAL
Contact: Simpson A.J.G.
COMMENT
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tbl=RC3&tbl2=RC3-PT0003-241199-011-d10&tbl3=1999-11-24&tbl4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 403.
FEATURES
Location/Qualifiers
1..403
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="PT0003"
/note="Organ: pnet; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 22; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 4.7;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCATTGG 22
 |||||
 Db 368 GATGCTGAAGTCGCTCATTGG 347

RESULT 2
 AA905101
 LOCUS ok05q06.s1 Soares NFL T.GBC.S1 Homo sapiens cDNA clone EST 09-JUN-1998
 DEFINITION IMAGE:1507354 3', mRNA sequence.

ACCESSION AA905101
 VERSION AA905101.1 GI:3040224
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 451)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 853 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 420.
 Location/Qualifiers
 1..451
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1507354"
 /lab_host="DH10B"
 /clone_lib="Soares NFL T.GBC.S1"
 /note="Organ: pooled; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES
 source
 1..451
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1507354"
 /lab_host="DH10B"
 /clone_lib="Soares NFL T.GBC.S1"
 /note="Organ: pooled; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 100.0%; Score 22; DB 1; Length 451;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCATTGG 22
 |||||
 Db 411 GATGCTGAAGTCGCTCATTGG 432

RESULT 3
 AW173368
 LOCUS x178f06.x1 Soares NFL T.GBC.S1 Homo sapiens cDNA clone EST 16-NOV-1999
 DEFINITION IMAGE:2663363 3', mRNA sequence.

ACCESSION AW173368
 VERSION AW173368.1 GI:6439316
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 493)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 471.
 Location/Qualifiers
 1..493
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2663363"
 /lab_host="DH10B"
 /clone_lib="Soares NFL T.GBC.S1"
 /note="Organ: pooled; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

FEATURES
 source
 1..493
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2663363"
 /lab_host="DH10B"
 /clone_lib="Soares NFL T.GBC.S1"
 /note="Organ: pooled; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 100.0%; Score 22; DB 2; Length 493;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGCTCATTGG 22
 |||||
 Db 467 GATGCTGAAGTCGCTCATTGG 488

RESULT 4
 AI804454
 LOCUS tc7le02.x1 Soares NHMPu_S1 Homo sapiens cDNA clone EST 13-DEC-1999
 DEFINITION 3', mRNA sequence.

ACCESSION AI804454
 VERSION AI804454.1 GI:5369926
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 514)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1553 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 469.
 Location/Qualifiers
 1..514
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2070074"

/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares_NhWMPu_S1"
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NhWMPu, and fetal heart NBHH19) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATGG 22
 |||||
 Db 453 GATGCTGAAGTCGTCCTCATGG 474

RESULT 5

AI082809
 LOCUS
 DEFINITION OX78a02.x1 Soares_NhWMPu_S1 Homo sapiens cDNA clone IMAGE:1662410
 3', mRNA sequence.

ACCESSION AI082809

VERSION AI082809.1 GI:3417785

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 531)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 902 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 451.

FEATURES

Location/Qualifiers

1..531
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1662410"
 /tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares_NhWMPu_S1"
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NhWMPu, and fetal heart NBHH19) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATGG 22
 |||||
 Db 453 GATGCTGAAGTCGTCCTCATGG 474

RESULT 7

BE466421
 LOCUS
 DEFINITION hz21e08.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3208646 3', mRNA sequence.

Query Match 100.0%; Score 22; DB 1; Length 531;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATGG 22
 |||||
 Db 453 GATGCTGAAGTCGTCCTCATGG 474

RESULT 6

AI768516
 LOCUS
 DEFINITION wh22g11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2381540 3', mRNA sequence.

ACCESSION AI768516

VERSION AI768516.1 GI:5235025

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 554)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 766 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 463.

Location/Qualifiers

1..554

FEATURES

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2381540"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Kid11"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI_CGAP_Kid3 was

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 1322376-1323911, 1456007-1456775, and

1500552-1502855). Subtraction by Bento Soares and M.

Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCCTCATGG 22
 |||||
 Db 446 GATGCTGAAGTCGTCCTCATGG 467

RESULT 7

BE466421
 LOCUS
 DEFINITION hz21e08.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3208646 3', mRNA sequence.

Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 654 Std Error: 0.00.
Location/Qualifiers
1. .654
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
/clone_lib="GRN_PREHEP"
/note="oligo dt primed, full-length enriched cDNA library from DMSO-treated HES cell line H9 (p22) maintained in feeder-free conditions"

FEATURES
source
Query Match 100.0%; Score 22; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 130 GATGCTGAAGTCGTCCTCATTGG 109
|||||

RESULT 9
BQ016066
LOCUS
DEFINITION
BQ016066
VERSION
BQ016066.1 GI:19751343
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 687)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5886984"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DT1"
/note="Organ: Lung; Vector: pTT3-Pac (Pharmacia) with a modified polylinker; Site: 1: Eco RI; Site 2: Not I; NCI-CGAP DT1 is a normalized cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer

FEATURES
source
Query Match 100.0%; Score 22; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 1 GATGCTGAAGTCGTCCTCATTGG 22
|||||
DB 447 GATGCTGAAGTCGTCCTCATTGG 468
|||||

RESULT 8
CN308828/c
LOCUS
DEFINITION
CN308828
ACCESSION
CN308828
VERSION
CN308828.1 GI:47325242
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 654)
Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guehler, K., Rao, M.S., Mandalan, R., Lebkowski, J and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R

FEATURES
source
Query Match 100.0%; Score 22; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mism

containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT733-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTCGG.

TAG_TISSUE=lung metatastic chondrosarcoma
TAG_LIB=UI-H-DT1
TAG_SEQ=AACTGTTCGG

ORIGIN

Query Match 100.0%; Score 22; DB 5; Length 687;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTGG 22
|||||
Db 467 GATGCTGAAGTCGTCTCATTGG 488

RESULT 10

BQ002692
LOCUS
DEFINITION
IMAGE:5843871 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=yes

FEATURES

source
1..782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5843871"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_E11"
/note="Organ: Left Pelvis; Vector: pT733-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI_CGAP_E11 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-805, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT733-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTCGG."

TAG_LIB=UI-H-E11
TAG_SEQ=ACACTTGCAC"

ORIGIN

Query Match 100.0%; Score 22; DB 5; Length 782;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTGG 22
|||||
Db 463 GATGCTGAAGTCGTCTCATTGG 484

RESULT 11

BG283778
LOCUS
DEFINITION
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10417 row: j column: 20
High quality sequence stop: 610.

FEATURES

Location/Qualifiers
1..895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4520179"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 91"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 22; DB 4; Length 895;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTGG 22
|||||
Db 296 GATGCTGAAGTCGTCTCATTGG 275

RESULT 12

BG786617
LOCUS
DEFINITION
mRNA sequence.

ACCESSION
VERSION
KEYWORDS

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 901)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: DCTD/DTF/Gazdar
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Incyte Genomics, Inc.
           Cloned through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LAM9650 row: j column: 12
           High quality sequence start: 10
           High quality sequence stop: 596.
           Location/Qualifiers
             1..901
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:3881579"
               /tissue_type="large cell carcinoma"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_68"
               /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
               Average insert size 1.8 kb. Library constructed by Life
               Technologies."
ORIGIN
  Query Match      100.0%; Score 22; DB 2; Length 901;
  Best Local Similarity 100.0%; Pred. No. 5.1;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGCTGAAGTCGTCTCATTTGG 22
    |||||
Db 167 GATGCTGAAGTCGTCTCATTTGG 146

RESULT 13
BG546951/c
LOCUS      BG546951
DEFINITION BG546951 901 bp mRNA linear EST 04-APR-2001
           mRNA sequence.
ACCESSION  BG546951
VERSION    BG546951.1 GI:13545616
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 901)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Cloned through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LAM10421 row: 1 column: 04
           High quality sequence stop: 766.
           Location/Qualifiers
             1..943
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:4521747"
               /tissue_type="embryonal carcinoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_92"
               /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
               Average insert size 2.5 kb. Library enriched for
               full-length clones and constructed by Life Technologies.
               Note: this is a NIH_MGC Library."
ORIGIN
  Query Match      100.0%; Score 22; DB 4; Length 943;
  Best Local Similarity 100.0%; Pred. No. 5.1;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 943)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Cloned through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LAM10421 row: 1 column: 04
           High quality sequence stop: 766.
           Location/Qualifiers
             1..943
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:4521747"
               /tissue_type="embryonal carcinoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH_MGC_92"
               /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
               Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
               Average insert size 2.5 kb. Library enriched for
               full-length clones and constructed by Life Technologies.
               Note: this is a NIH_MGC Library."
ORIGIN
  Query Match      100.0%; Score 22; DB 4; Length 943;
  Best Local Similarity 100.0%; Pred. No. 5.1;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GATCGTGAAGTCGTCCTCATTTGG 22
 |||||
 Db 394 GATCGTGAAGTCGTCCTCATTTGG 373

RESULT 15
 HSM801459/c
 LOCUS Homo sapiens mRNA; cDNA DKFZp434G0715 (from clone DKFZp434G0715).
 DEFINITION Homo sapiens mRNA; cDNA DKFZp434G0715 (from clone DKFZp434G0715).
 ACCESSION AL133605
 VERSION AL133605.1 GI:6599213
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4308)
 Bloembergen, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,
 Oanger, A., Fobo, G., Han, M. and Wiemann, S.
 The German cDNA Consortium
 Direct Submission
 Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp434G0715) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp434G0715
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/.
 FEATURES
 source
 1..4308
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp434G0715"
 /tissue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /note="pallino 2, N-terminus truncated"
 gene
 1..4308
 /gene="DKFZp434G0715"
 CDS
 <1..150
 /codon_start=1
 /gene="DKFZp434G0715"
 /product="hypothetical protein"
 /protein_id="CAH56388.1"
 /db_xref="GI:52545575"
 /translation="VCSEKSAKYWSQIPLPHGTHAFHAACPFCATQLVGEQNCIKLIF
 QGPID"

ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 4308;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCGTGAAGTCGTCCTCATTTGG 22
 |||||
 Db 3846 GATCGTGAAGTCGTCCTCATTTGG 3825

Search completed: July 20, 2005, 18:56:21
 Job time : 1582.48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 15:43:06 ; Search time 1716.52 Seconds
(without alignments)
532.206 Million cell updates/sec

Title: US-10-041-030-8

Perfect score: 24

Sequence: 1 ccagtagtttagccttggtgctt 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:.*
2: gb_est2:.*
3: gb_hic:.*
4: gb_est3:.*
5: gb_est4:.*
6: gb_est5:.*
7: gb_est6:.*
8: gb_gsel:.*
9: gb_gsel2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24	100.0	214	AW272021	AW272021 xr45b04.x
C 2	24	100.0	295	BI038270	BI038270 RC5-NT026
C 3	24	100.0	531	AI082809	AI082809 ox78a02.x
C 4	24	100.0	554	AI768516	AI768516 wh22g11.x
C 5	24	100.0	654	CN308828	CN308828 170006001
C 6	24	100.0	687	BQ016066	BQ016066 UI-H-DTI-
C 7	24	100.0	782	BQ002692	BQ002692 UI-H-E11-
C 8	24	100.0	895	EG283778	EG283778 602408077
C 9	24	100.0	901	EG546951	EG546951 602573943
C 10	24	100.0	943	EG388221	EG388221 602413214
C 11	24	100.0	4308	HSW801459	HSW801459 Homo sapi
C 12	23	95.8	555	BE466421	BE466421 h221e08.x
C 13	22.4	93.3	274	BQ091930	BQ091930 UNM17E05
C 14	22.4	93.3	403	AW386118	AW386118 RC3-PT000
C 15	20.8	86.7	171	CB117270	CB117270 K-EST0162
C 16	19.2	80.0	101	N99024	N99024 za67f06.r1
C 17	19.2	80.0	173	BI001452	BI001452 MR3-HN012
C 18	19.2	80.0	228	BI001750	BI001750 MR3-HN014
C 19	19.2	80.0	243	BI001761	BI001761 MR3-HN014
C 20	19.2	80.0	265	BQ359661	BQ359661 MR3-HN014
C 21	19.2	80.0	332	AA617629	AA617629 np34g08.s
C 22	19.2	80.0	462	AI822085	AI822085 za92h12.y
C 23	19.2	80.0	492	BM996453	BM996453 UI-H-DT0-
C 24	19.2	80.0	634	BM970838	BM970838 UI-CF-EC1

C 25	19.2	80.0	635	2	BB578089	BB578089
C 26	19.2	80.0	725	6	CA502940	CA502940 UI-CF-PN0
C 27	19.2	80.0	777	5	BX443706	BX443706 BX443706
C 28	19.2	80.0	919	5	BX372042	BX372042 BX372042
C 29	19.2	80.0	1692	3	CR600885	CR600885 full-leng
C 30	19.2	80.0	2280	3	CR604190	CR604190 full-leng
C 31	18.8	78.3	198	5	BQ095095	BQ095095 Kk12h10.y
C 32	18.8	78.3	437	5	BY001426	BY001426 BY001426
C 33	18.8	78.3	591	9	CR189890	CR189890 Reverse s
C 34	18.8	78.3	644	9	CE192848	CE192848 t1gr-988-
C 35	18.8	78.3	645	2	BE911617	BE911617 601663094
C 36	18.8	78.3	1180	5	BX334080	BX334080 BX334080
C 37	18.4	76.7	539	7	CO585167	CO585167 DG2-1111
C 38	18.4	76.7	562	7	CO586470	CO586470 DG2-131a6
C 39	18.4	76.7	576	7	CO716806	CO716806 DG14-30k2
C 40	18.4	76.7	579	7	CO699795	CO699795 DG32-166b
C 41	18.4	76.7	600	7	CO696308	CO696308 DG11-96n1
C 42	18.4	76.7	628	7	CO591703	CO591703 DG2-53f18
C 43	18.4	76.7	631	7	CO588078	CO588078 DG2-19d14
C 44	18.4	76.7	637	7	CO588492	CO588492 DG2-2115
C 45	18.4	76.7	640	7	CO693480	CO693480 DG11-63a1

ALIGNMENTS

RESULT 1
AW272021/c
LOCUS
DEFINITION
xr45b04.x1 NCI CGAP_Ov26 Homo sapiens cDNA clone IMAGE:2763055 3',
mRNA;sequence.
ACCESSION
AW272021
VERSION
AW272021.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 214)
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Monica Brown, M.D., Elise Kohn, M.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES

1..214
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2763055"
/sex="female"
/tissue_type="papillary serous carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pAMP1; mRNA made from
papillary serous ovarian carcinoma, cDNA made by oligo-dT
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 500 bp. Primary library,
non-amplified."

ORIGIN

Query Match 100.0%; Score 24; DB 2; Length 214;

Best Local Similarity 100.0%; Pred. No. 1.5; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24
Db 57 CCAGTAGTTAGCCTTTGGCTT 34

RESULT 2
BI038270/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

295 bp mRNA linear EST 14-JUN-2001
RC5-NT0266-180101-011-B12 NT0266 Homo sapiens cDNA, mRNA sequence.
BI038270
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
2002663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&t2=RC5-NT0266-
180101-011-B12&t3=2001-01-18&t4=1)

Seq primer: puc 18 forward
High quality sequence start: 28
High quality sequence stop: 294.
Location/Qualifiers
1. .295
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0266"

/notes="Organ: nervous tumor; Vector: puc18; Site 1: Sma1;
Site 2: Sma1; A mini-library was made by cloning products
derived from ORFEST PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 100.0%; Score 24; DB 4; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTAGCCTTTGGCTT 24
Db 60 CCAGTAGTTAGCCTTTGGCTT 37

RESULT 3
AI082809/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

554 bp mRNA linear EST 20-DEC-1999
wh22g11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2381540 3',
mRNA sequence.
AI768516
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

531 bp mRNA linear EST 24-SEP-1998
ox78a02.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662410
3', mRNA sequence.
AI082809
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

531 bp mRNA linear EST 24-SEP-1998
ox78a02.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662410
3', mRNA sequence.
AI082809
EST.
Homo sapiens (human)

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

531 bp mRNA linear EST 24-SEP-1998
ox78a02.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662410
3', mRNA sequence.
AI082809
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 902 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 451.
Location/Qualifiers
1. .531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:1662410"

/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares_NhMpu_S1"
/note="Organ: mixed (see below); Vector: p77T3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

ORIGIN
Query Match 100.0%; Score 24; DB 1; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
Db 526 CCAGTAGTTAGCCTTTGTGGCTT 503

RESULT 4
AI768516/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

554 bp mRNA linear EST 20-DEC-1999
wh22g11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2381540 3',
mRNA sequence.
AI768516
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

531 bp mRNA linear EST 24-SEP-1998
ox78a02.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662410
3', mRNA sequence.
AI082809
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

531 bp mRNA linear EST 24-SEP-1998
ox78a02.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662410
3', mRNA sequence.
AI082809
EST.
Homo sapiens (human)

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 766 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 463.
 Location/Qualifiers
 1. .554
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2381540"
 /lab_hosts="DH10B"
 /clone_lib="NCI CGAP Kid11"
 /notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 100.0%; Score 24; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAGTAGTTAGCCTTTGGCTT 24

|||||
 Db 519 CCAGTAGTTAGCCTTTGGCTT 496

RESULT 5

CN308828 1700600172033 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence. EST 16-MAY-2004
 LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Brandenberger R
 Regenerative Medicine
 Genon Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@genon.com
 Insert Length: 654 Std Error: 0.00.
 Location/Qualifiers
 1. .654
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, DMSO-treated H9 cell line"

FEATURES

source

Query Match 100.0%; Score 24; DB 5; Length 687;
 Best Local Similarity 100.0%; Pred. No. 1.8;

ORIGIN

Query Match 100.0%; Score 24; DB 7; Length 654;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CCAGTAGTTAGCCTTTGGCTT 24
 |||||
 Db 57 CCAGTAGTTAGCCTTTGGCTT 80

RESULT 6

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ016066 687 bp mRNA linear EST 17-JUN-2002
 UI-H-DTI-awb-a-01-0-UI.81 NCI CGAP_DTI Homo sapiens cDNA clone
 IMAGE:5886984 3', mRNA sequence.
 BQ016066
 BQ016066.1 GI:19751343
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 687)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1. .687
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5886984"
 /tissue_type="Metastatic Chondrosarcoma"
 /dev_host="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP DT1"
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP DT1 is a normalized cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTCCG.
 TAG TISSUE=lung metastatic chondrosarcoma
 TAG_LIB=UI-H-DTI
 TAG_SEQ=AACTGTTCCG"

ORIGIN

Query Match 100.0%; Score 24; DB 5; Length 687;
 Best Local Similarity 100.0%; Pred. No. 1.8;

/clone lib="GRN PREHEP"
 /note="oligo dT primed, full-length enriched cDNA library from DMSO-treated HES cell line H9 (p22) maintained in feeder-free conditions"

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
 Db 540 CCAGTAGTTAGCCTTTGTGGCTT 517

RESULT 7
 BQ002692/c
 LOCUS
 DEFINITION
 UI-H-E11-ayu-1-16-0-UI.s1 NCI CGAP_E11 Homo sapiens cDNA clone
 IMAGE:5843871 3', mRNA sequence.
 ACCESSION
 VERSION
 BQ002692
 BQ002692.1 GI:19727592
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 782)
 REFERENCE
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS
 TITLE
 Tumor Gene Index
 JOURNAL
 Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 source
 1..782
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5843871"
 /tissue_type="Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_E11"
 /notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
 with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP E11 is a normalized cDNA library containing the
 following tissue(s): Chondrosarcoma. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 ACACCTTGAC.
 TAG TISSUE=chondrosarcoma
 TAG LIB=UI-H-E11
 TAG_SEQ=ACACTTGAC"

ORIGIN
 Query Match 100.0%; Score 24; DB 5; Length 782;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
 Db 536 CCAGTAGTTAGCCTTTGTGGCTT 513

RESULT 8

BG283778
 LOCUS
 DEFINITION
 602408077F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520179 5',
 mRNA sequence.
 ACCESSION
 VERSION
 BG283778
 BG283778.1 GI:13034060
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 895)
 REFERENCE
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10417 row: j column: 20
 High quality sequence stop: 610.

FEATURES
 source
 1..895
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520179"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 100.0%; Score 24; DB 4; Length 895;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
 Db 223 CCAGTAGTTAGCCTTTGTGGCTT 246

RESULT 9
 BG546951
 LOCUS
 DEFINITION
 602573943F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4702067 5',
 mRNA sequence.
 ACCESSION
 VERSION
 BG546951
 BG546951.1 GI:13545616
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 901)
 REFERENCE
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

ORIGIN
 Query Match 100.0%; Score 24; DB 5; Length 782;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCAGTAGTTAGCCTTTGTGGCTT 24
 Db 536 CCAGTAGTTAGCCTTTGTGGCTT 513

RESULT 8

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCMI538 row: e column: 12
High quality sequence stop: 751.
Location/Qualifiers

FEATURES

source

1. .901
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4702067"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 77"
/notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 Kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 24; DB 4; Length 901;
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTAGCCTTTGTGGCTT 24
|||||
DB 153 CCAGTAGTTTAGCCTTTGTGGCTT 176

RESULT 10

BG388221

LOCUS 602413214F1 NIH_MGC_92 943 bp mRNA linear EST 12-MAR-2001
DEFINITION mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCMI538 row: 1 column: 04

High quality sequence stop: 766.
Location/Qualifiers

FEATURES

source

1. .943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4521747"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 24; DB 4; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTAGCCTTTGTGGCTT 24
|||||
DB 321 CCAGTAGTTTAGCCTTTGTGGCTT 344

RESULT 11

HSM801459

LOCUS HSM801459 4308 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp434G0715 (from clone DKFZp434G0715).
ACCESSION AL133605

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

CONSRMT

TITLE

JOURNAL

COMMENT

1 (bases 1 to 4308)
Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,
Oanger, A., Fobbo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr. 1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.

This clone (DKFZp434G0715) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp434G0715>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

source

1. .4308
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434G0715"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stages="adult"
/note="pellino 2, N-terminus truncated"
1. .4308
/gene="DKFZp434G0715"
1. .150
/gene="DKFZp434G0715"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH56388.1"
/db_xref="GI:52545575"
/translation="VCSEKSAKYWSQIPLPHGTHAFHAACPFCATQLVGEQNCIKLIF
QGPID"

ORIGIN

Query Match 100.0%; Score 24; DB 3; Length 4308;
Best Local Similarity 100.0%; Pred. No. 2.4; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTAGCCTTTGTGGCTT 24
|||||
DB 3773 CCAGTAGTTTAGCCTTTGTGGCTT 3796

RESULT 12
 BE466421/c
 LOCUS
 DEFINITION hz21e08.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3208646 3', mRNA linear EST 27-JUL-2000
 ACCESSION BE466421
 VERSION BE466421
 KEYWORDS BE466421.1 GI:9512283
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 555)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seq primer: -40UP from Gibco
 High quality sequence stop: 453.
 Location/Qualifiers
 1..555

FEATURES
source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3208646"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_GC6"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIds 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 95.8%; Score 23; DB 2; Length 555;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CAGTAGTTTAGCCTTTGTGGCTT 24
 |||||
 Db 519 CAGTAGTTTAGCCTTTGTGGCTT 497

RESULT 13
 BQ091930/c
 LOCUS
 DEFINITION UMN17E05 Canine Brain cDNA Library Canis familiaris cDNA 5', mRNA linear EST 08-APR-2002
 ACCESSION BQ091930
 VERSION BQ091930.1 GI:20072538
 KEYWORDS EST.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 1 (bases 1 to 274)

AUTHORS

TITLE
JOURNAL
COMMENT

Roberts, M.C., Hendrickson, J.A., Hoffmann, D.E., Flickinger, G.H., Rutherford, M.S. and Mickelson, J.R.
 University of Minnesota Canine Brain EST Project Unpublished (2001)
 Contact: Mickelson, J.R.
 Veterinary Pathobiology
 University of Minnesota
 1988 Fitch Avenue, University of Minnesota, St. Paul, MN 55108, USA
 Tel: 612 624 1246
 Fax: 612 625 0204
 Email: mick001@umn.edu
 Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
 1..274
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /sex="Male"
 /clone_lib="Canine Brain cDNA Library"
 /note="Organ: Brain; Vector: pSPORT1 (Gibco BRL); Site_1: NotI; Site_2: SalI; Tissue was taken from the frontal, occipital, temporal and parietal lobes, olfactory bulb, hippocampus, cerebellum, thalamus, hypothalamus, midbrain, pons, and medulla."

ORIGIN

Query Match 93.3%; Score 22.4; DB 5; Length 274;
 Best Local Similarity 95.8%; Pred. No. 8.8;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCAGTAGTTTAGCCTTTGTGGCTT 24
 |||||
 Db 189 CCAGTAGTTTAGCCTTTGTGGCTT 166

RESULT 14

AW386118
 LOCUS
 DEFINITION RC3-PT0003-241199-011-d10 PT0003 Homo sapiens cDNA, mRNA linear EST 04-FEB-2000
 ACCESSION AW386118
 VERSION AW386118.1 GI:6890777
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 403)
 REFERENCE HCGP <http://www.ludwig.org.br/ORESTES>.
 AUTHORS The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-PT0003-241199-011-d10&t3=1999-11-24&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 403.
 Location/Qualifiers
 1..403

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="PT0003"
 /note="Organ: pnet; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 93.3%; Score 22.4; DB 2; Length 403;
Best Local Similarity 95.8%; Pred. No. 9.4;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCAGTAGTTTAGCCTTTGGGCTT 24
| | | | | | | | | | | | | | | | | | | | | |
Db 295 CCAGTAGTTTAGCCTTTGGGCTT 318

RESULT 15

CB117270
LOCUS K-EST0162426 L4SNU368 Homo sapiens CDNA clone L4SNU368-26-E03 5',
DEFINITION mRNA sequence.
ACCESSION CB117270
VERSION CB117270.1 GI:27943077
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 171)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished. (2002)

COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 26 row: E column: 03
High quality sequence stop: 171.
Location/Qualifiers
1. .171

FEATURES

source
1. .171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L4SNU368-26-E03"
/sex="M"
/tissue_type="Liver"
/cell_type="Polygonal"
/cell_line="SNU-368"
/lab_host="Top10F"
/clone_lib="L4SNU368"
/notes="Organ: Liver; Vector: pcNS-D2; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 86.7%; Score 20.8; DB 6; Length 171;
Best Local Similarity 91.7%; Pred. No. 47;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCAGTAGTTTAGCCTTTGGGCTT 24
| | | | | | | | | | | | | | | | | | | | | |
Db 69 CCACTAGTTTATCCCTTTGGGCTT 92

Search completed: July 20, 2005, 18:56:26
Job time : 1721.52 secs

This Page Blank (uspto)